

Query Match	99.8%;	Score 899;	DB 6;	Length 3294;
Best Local Similarity	100.0%;	Pred. No. 3.2e-154;		
Matches 901;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	TGTCGGAAATTACCGACCAACTTTGGTCGGTCAATTTAAATTCAAAAAATATTGTAAA	60	
Db	2255	TGTCGGAAATTACCGACCAACTTTGGTCGGTCAATTTAAATTCAAAAAATATTGTAAA	231	
QY	61	AAAAAACCGACCAAGTTGATTCGGTATTTTAAATTATGTAATAAAAGATTCCTACTTCGG	120	
Db	2315	AAAAAACCGACCAAGTTGATTCGGTATTTTAAATTATGTAATAAAAGATTCCTACTTCGG	237	
QY	121	GAATCGAACCCGGGTCTGTACTATGCGAAGATCACTATTACCACTAGACCAATTGGTTCA	180	
Db	2375	GAATCGAACCCGGGTCTGTACTATGCGAAGATCACTATTACCACTAGACCAATTGGTTCA	243	
QY	181	TTTTTCTTTTAAAGCTGCTCTTTATTTCAATTATTACTCTCTTAATTATATATTTTTTGCACGAA	240	

LOCUS AX068013 4312 bp DNA linear PAT 19-JAN-2001
DEFINITION Sequence 2 from Patent WO007187.
ACCESSION AX068013
VERSION AX068013.1 GI:12329821
KEYWORDS Nicotiana tabacum (common tobacco)
SOURCE Nicotiana tabacum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 Roitsch, T.D.
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 007187-A 2 21-DEC-2000;
Roitsch, Thomas, Dr. (DE)
FEATURES
source
1. 4312
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Query Match 99.8%; Score 899; DB 6; Length 4312;
Best Local Similarity 99.8%; Pred. No. 3.1e-154;
Matches 899; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TGGTCGGAATATCCGACCAACTTGGTCGGTCAATTAATTCAAAAAATATTTGATA 60
DB 3273 TGGTCGGAATATCCGACCAACTTGGTCGGTCAATTAATTCAAAAAATATTTGATA 3332
QY 61 AAAAAACCGACCAAGTTGATCGGTATTTAAATATGTAATAAAAAAGATTCACTATCTGG 120
DB 3333 AAAAAACCGACCAAGTTGATCGGTATTTAAATATGTAATAAAAAAGATTCACTATCTGG 3392
QY 121 GAATCGAACCGGGTCGTACTATGGAAGATCTATTCACATAGACCAATGGTTCA 180
DB 3393 GAATCGAACCGGGTCGTACTATGGAAGATCTATTCACATAGACCAATGGTTCA 3452
QY 181 TTTTGTATTAAGACTGCTTTTATTTGATTTATCTTTTAAATATATATTTTTCACGAAA 240
DB 3453 TTTTGTATTAAGACTGCTTTTATTTGATTTATCTTTTAAATATATATTTTTCACGAAA 3512
QY 241 ATACCGACCAAGTTGGTCGATTTATTTAAAGTAATAATTTACTTACCAAGTTGGTCG 300
DB 3513 ATACCGACCAAGTTGGTCGATTTATTTAAAGTAATAATTTACTTACCAAGTTGGTCG 3572
QY 301 ATTTTATTAATGATCGCGGATTAACCGACCAATTTGGTAGGTTTATTAATTTAA 360
DB 3573 ATTTTATTAATGATCGCGGATTAACCGACCAATTTGGTAGGTTTATTAATTTAA 3632
QY 361 TTTTATTTATTTTAAATGAAAAAATAACCAAAAGTTAGTCGGTTCTTGAACATAAAAT 420
DB 3633 TTTTATTTATTTTAAATGAAAAAATAACCAAAAGTTAGTCGGTTCTTGAACATAAAAT 3692
QY 421 TCGGGGACTCAAAATAGTTTCCCGATTTTCCGCAAGAAAAACCGACCAAGTTGG 480
DB 3693 TCGGGGACTCAAAATAGTTTCCCGATTTTCCGCAAGAAAAACCGACCAAGTTGG 3752
QY 481 TCGGTTTCGTAATAAAAAAATTTAAATAATATTTTAAACCGACCAACTTTA 540
DB 3753 TCGGTTTCGTAATAAAAAAATTTAAATAATATTTTAAACCGACCAACTTTA 3812
QY 541 GTCGGTTTTTGGTCGATTTTGTGACCGACCAAGTTGGTCGGTTCGACCTTGGTCGGTTT 600
DB 3813 GTCGGTTTTTGGTCGATTTTGTGACCGACCAAGTTGGTCGGTTCGACCTTGGTCGGTTT 3872
QY 601 TTGCGGAATTTCTAGTAGTACCGACCTGTAGCTTCGGGAGAAATTTGTATATGA 660
DB 3873 TTGCGGAATTTCTAGTAGTACCGACCTGTAGCTTCGGGAGAAATTTGTATATGA 3932
QY 661 TATGTGTATATCTTAAATAGTTTAAATTAAGAAACGNNCCACCTGGAATCTAGAGGC 720
DB 3933 TATGTGTATATCTTAAATAGTTTAAATTAAGAAACGNNCCACCTGGAATCTAGAGGC 3992

QY 721 TTTAGGGCAGCTAGATGAGAGATAAAGTGTCTTCGTCGGTAAAAATATCTTGATCCG 780
DB 3993 TTTAGGGCAGCTAGATGAGAGATAAAGTGTCTTCGTCGGTAAAAATATCTTGATCCG 4052
QY 781 CCTATGATGTAAGTACTTCTCGPCTTAATCAGAGGTTTCGACTTCGAGCTCCAGATA 840
DB 4053 CCTATGATGTAAGTACTTCTCGPCTTAATCAGAGGTTTCGACTTCGAGCTCCAGATA 4112
QY 841 TAAACTATAGACTCGTCTTTTATAGCACCTTTTAAATAGACTATGACTTCATCTGATTCT 900
DB 4113 TAAACTATAGACTCGTCTTTTATAGCACCTTTTAAATAGACTATGACTTCATCTGATTCT 4172
QY 901 C 901
DB 4173 C 4173
RESULT 4
LOCUS AF376772S1 3293 bp DNA linear PLN 29-MAY-2001
DEFINITION Nicotiana tabacum extracellular invertase Nin88 (Nin88) gene, promoter and partial cds.
ACCESSION AF376772
VERSION AF376772.1 GI:14211752
KEYWORDS
SEGMENT 1 of 2
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Nicotiana.
AUTHORS Goetz, M., Godt, D.E., Guivarc'h, A., Kahmann, U., Chriqui, D. and Roitsch, T.
TITLE Induction of male sterility in plants by metabolic engineering of the carbohydrate supply
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (2001) In press
REFERENCE 2 (bases 1 to 3293)
AUTHORS Goetz, M., Godt, D.E. and Roitsch, T.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-2001) Plant Physiology and Cell Biology, University of Regensburg, Universitaetstr. 31, Regensburg 93053, Germany
FEATURES
source
1. 3293
/organism="Nicotiana tabacum"
/mol_type="genomic DNA"
/db_xref="taxon:4097"
promoter
1. 3290
/gene="Nin88"
/gene="Nin88"
/product="extracellular invertase Nin88"
3291..>3293
/codon_start=1
/gene="Nin88"
/product="extracellular invertase Nin88"
/protein_id="AAK57503.1"
/db_xref="GI:14211754"
/translation="M"
ORIGIN
Query Match 98.1%; Score 883.8; DB 8; Length 3293;
Best Local Similarity 99.4%; Pred. No. 1.9e-151;
Matches 896; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 TGGTCGGAATATCCGACCAACTTGGTCGGTCAATTAATTCAAAAAATATTTGATAA 60
DB 2255 TGGTCGGAATATCCGACCAACTTGGTCGGTCAATTAATTCAAAAAATATTTGATAA 2314
QY 61 AAAAAACCGACCAAGTTGATCGGTATTTAAATATGTAATAAAAAAGATTTCACATCTCGG 120

two late


```

/mol_type="genomic DNA"
/db_xref="taxon:4096"
/clone="lambda NSCbb1"
/tissue_type="leaf"
/clone_in="lambda DASH genomic library"
complement(1870..2673)
/gene="Lhcb1*6"
complement(1870..2673)
/gene="Lhcb1*6"
/function="antenna complex"
/codon_start=1
/product="light harvesting chlorophyll a/b-binding protein"
/protein_id="BAA25392.1"
/db_xref="GI:3036948"
/translations="MAASTMALSSSFAGKAVKLSPPSSSEITGNGKVTMRKTSKAKP
VSSGWPYGPDRVKYLPFSGESPSYLTGFPGDYGMDTAGLSADPETPAKKNRELEVI
HCRWMLGALGVCPPELLARNGVKFGEAVWFKAGSOIPSEGGLDYLGNSPLVHAQSIL
AIWACQVLMGAVEGYRVAGGPLGEVVDPLYPGGSFDP/LGLAEDPEAFALVKVEIKN
GLRMFSMFGFVQALVTGKPLENLADHLADPVNNNAWAYATNFVPGK"
complement(2572..2673)
/gene="Lhcb1*6"
/notes="putative"
complement(1873..2571)
/gene="Lhcb1*6"
/product="light harvesting chlorophyll a/b-binding protein"
/notes="putative"
complement(2760..2767)
/notes="Lhcb1*6
putative"
4272..4278
/notes="Lhcb1*5
putative"
4363..5163
/gene="Lhcb1*5"
4363..5163
/gene="Lhcb1*5"
/function="antenna complex"
/codon_start=1
/product="light harvesting chlorophyll a/b-binding protein"
/protein_id="BAA25393.1"
/db_xref="GI:3036949"
/translations="MAASTMALSSSFAGKAVKLSPPSSSEITGNGKVTMRKTSKAKPV
SSSLWYGPDRVKYLPFSGESPSYLTGFPGDYGMDTAGLSADPETPAKKNRELEVIH
CRWMLGALGVCPPELLARNGVKFGEAVWFKAGSOIPSEGGLDYLGNSPLVHAQSILA
IWACQVLMGAVEGYRVAGGPLGEVVDPLYPGGSFDP/LGLAEDPEAFALVKVEIKNG
RLRMFSMFGFVQALVTGKPLENLADHLADPVNNNAWAYATNFVPGK"
4363..4461
/gene="Lhcb1*5"
/notes="putative"
4462..5160
/gene="Lhcb1*5"
/product="light harvesting chlorophyll a/b-binding protein"
/notes="putative"

ORIGIN
Query Match 8.5%; Score 76.4; DB 8; Length 7299;
Best Local Similarity 79.5%; Pred. No. 0.00024;
Matches 89; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 628 CCTGTAGCTCGGGAGAAATTTGTATATGTATGTATGTATATCTTAAATGATTAT 687
Db 3928 CCGTAACCTCAGCAAACTCTGTATATGTATGTATGTATATCTTAAATGATTAT 3869

QY 688 TTAAGAACGNGCACCTCGAATACTAGAACCTTTAGGGGCGCACTAGATGAG 739
Db 3868 TTAATCAGCTGACACCTTAATGCTAAAGCCCTTAGGGGCGCAGATCAG 3817

```

```

AF376772S1/c AF376772S1 3293 bp DNA linear PLN 29-MAY-2001
LOCUS Nicotiana tabacum extracellular invertase Nin88 (Nin88) gene,
DEFINITION promoter and partial cds.
ACCESSION AF376772
VERSION AF376772.1 GI:14211752
KEYWORDS i of 2
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 3293)
AUTHORS Goetz,M., Godt,D.E., Guivarc'h,A., Kahmann,U., Chriqui,D. and
Roitsch,T.
TITLE Induction of male sterility in plants by metabolic engineering of
the carbohydrate supply
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (2001) In press
REFERENCE 2 (bases 1 to 3293)
AUTHORS Goetz,M., Godt,D.E. and Roitsch,T.
TITLE Direct Submision
JOURNAL Submitted (04-MAY-2001) Plant Physiology and Cell Biology,
University of Regensburg, Universitaetsstr. 31, Regensburg 93053,
Germany
FEATURES
Location/Qualifiers
1..3293
/organism="Nicotiana tabacum"
/mol_type="genomic DNA"
/db_xref="taxon:4097"
1..3290
/gene="Nin88"
<3291..>3293
/gene="Nin88"
/product="extracellular invertase Nin88"
3291..>3293
/gene="Nin88"
/codon_start=1
/product="extracellular invertase Nin88"
/protein_id="AAK57503.1"
/db_xref="GI:14211754"
/translation="M"

ORIGIN
Query Match 8.0%; Score 72.4; DB 8; Length 3293;
Best Local Similarity 60.3%; Pred. No. 0.0014;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCCGCCGAATTAACCGACCAATTTTGGTAGTGTATTAATTTTAA 366
Db 2819 TCAGAAATTCGACCAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760

QY 367 ---TTTATTTTAAATGAAAACCTACCAAGTTAGTCGGTTCTTGAACATAATTTTCG 423
Db 2759 ATTTTATTTTATTTTACGAAACCGACCACTTTGTCGGTTTTCTTTGGCGCAAAATGC 2700

QY 424 CGGGACTCAAAAATAGTTTCCCGCATTTTTCGGCAAGAAAACCGACCAAGTTGGTCG 483
Db 2699 GCGAACTATTTTTCAGTCCCGCGAAATTTATGTTTCAAGAACCGACTAATTGGTTA 2640

QY 484 GTTTCGTAAAAAATAAATAATTTAAAAATAATATTTTAAAAAACCGACCACTTAGTC 543
Db 2639 GTTTTCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2584

QY 544 GGTTTTGGTCGATTTTGTGA 565
Db 2583 GGTAAATTCGGCGGATCAATTA 2562

RESULT 8
AX068012/c AX068012 3294 bp DNA linear PAT 19-JAN-2001
DEFINITION Sequence 1 from Patent WO0077187.

```

ACCESSION AX068012
VERSION AX068012.1 GI:12329820
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 007187-A 1 21-DEC-2000;
Roitsch, Thomas, Dr. (DE)
FEATURES
source 1..3294
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Query Match 8.0%; Score 72.4; DB 6; Length 3294;
Best Local Similarity 60.3%; Pred. No. 0.0014;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;
QY 308 TAAATGATCCCGCAATTAACCGACCAATTTGGTAGGTTTTTAAATATATATTTTA- 366
DB 2819 TCAGAAATCGACCAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760
QY 367 ---TTTATTAAATGAAAAAATAACCAAGTTAGTCGGTTTCTTGAACATAAATTCG 423
DB 2759 ATTTTATTTTACGAAACCGCAACTTTGGTCGGTTTTTCTTTGGCGCAAAATGC 2700
QY 424 CGGACTCAAAATAGTTTCGGCATTTTGGCCCAAGAAACCGACCAAGTTGGTCG 483
DB 2699 GGGAAACTATTTTGAGTCCGCGAAATTTATGTTCAAGAAACCGACTAACTTTG 2640
QY 484 GTTTCGTAAAAAATAAATTTAAAAATATATTTAAAAAACCACCAACTTTAGTC 543
DB 2639 GTTTTCAATTAATAAATAA---TAAATTAATTAATAAARACCTACCAAAATGGTC 2584
QY 544 GGTTTTGGTCGATTTTGA 565
DB 2583 GGTAAATTCGGCGGATCATTTA 2562
RESULT 9
LOCUS AX068019/c 4135 bp DNA linear PAT 19-JAN-2001
DEFINITION Sequence 8 from Patent WO007187.
ACCESSION AX068019
VERSION AX068019.1 GI:12329827
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 007187-A 8 21-DEC-2000;
Roitsch, Thomas, Dr. (DE)
FEATURES
source 1..4135
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Fusion aus Promotor und codierender Sequenz in antisense-Orientierung"
ORIGIN
Query Match 8.0%; Score 72.4; DB 6; Length 4135;
Best Local Similarity 60.3%; Pred. No. 0.0014;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCCCGCAATTAACCGACCAATTTGGTAGGTTTTTAAATATATTTTA- 366
DB 2819 TCAGAAATCGACCAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760
QY 367 ---TTTATTAAATGAAAAAATAACCAAGTTAGTCGGTTTCTTGAACATAAATTCG 423
DB 2759 ATTTTATTTTACGAAACCGCAACTTTGGTCGGTTTTTCTTTGGCGCAAAATGC 2700
QY 424 CGGACTCAAAATAGTTTCGGCATTTTGGCCCAAGAAACCGACCAAGTTGGTCG 483
DB 2699 GGGAAACTATTTTGAGTCCGCGAAATTTATGTTCAAGAAACCGACTAACTTTG 2640
QY 484 GTTTCGTAAAAAATAAATTTAAAAATATATTTAAAAAACCACCAACTTTAGTC 543
DB 2639 GTTTTCAATTAATAAATAA---TAAATTAATTAATAAARACCTACCAAAATGGTC 2584
QY 544 GGTTTTGGTCGATTTTGA 565
DB 2583 GGTAAATTCGGCGGATCATTTA 2562
RESULT 10
LOCUS AX068013/c 4312 bp DNA linear PAT 19-JAN-2001
DEFINITION Sequence 2 from Patent WO007187.
ACCESSION AX068013
VERSION AX068013.1 GI:12329821
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 007187-A 2 21-DEC-2000;
Roitsch, Thomas, Dr. (DE)
FEATURES
Location/Qualifiers
source 1..4312
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Query Match 8.0%; Score 72.4; DB 6; Length 4312;
Best Local Similarity 60.3%; Pred. No. 0.0014;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;
QY 308 TAAATGATCCCGCAATTAACCGACCAATTTGGTAGGTTTTTAAATATATTTTA- 366
DB 3837 TCAGAAATCGACCAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 3778
QY 367 ---TTTATTAAATGAAAAAATAACCAAGTTAGTCGGTTTCTTGAACATAAATTCG 423
DB 3777 ATTTTATTTTACGAAACCGCAACTTTGGTCGGTTTTTCTTTGGCGCAAAATGC 3718
QY 424 CGGACTCAAAATAGTTTCGGCATTTTGGCCCAAGAAACCGACCAAGTTGGTCG 483
DB 3717 GGGAAACTATTTTGAGTCCGCGAAATTTATGTTTCAAGAAACCGACTAACTTTG 3658
QY 484 GTTTCGTAAAAAATAAATTTAAAAATATATTTAAAAAACCACCAACTTTAGTC 543
DB 3657 GTTTTCAATTAATAAATAA---TAAATTAATTAATAAARACCTACCAAAATGGTC 3602
QY 544 GGTTTTGGTCGATTTTGA 565
DB 3601 GGTAAATTCGGCGGATCATTTA 3580
RESULT 11
LOCUS AC090485/c 159636 bp DNA linear PLN 06-SEP-2001

```

DEFINITION Genomic Sequence for Oryza sativa, Nipponbare strain, clone
ACCESSION OSJNBa0067N01, from chromosome 3, complete sequence.
VERSION AC090485
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 159636)
AUTHORS de la Bastide,M., Spiegel,L., Preston,R., Kirchoff,K., King,L.,
Nascimento,L., Baker,J., Vil,M.D., Zutavern,T., Santos,L.,
Miller,B., Kuit,K., Cunnius,D.M., Rodriguez,S., Balija,V., Shah,R.,
Bahret,A., Yang,C., Bell,M., Palmer,L., O'Shaughnessy,A., Dedhia,N.
and McCombie,W.R.
TITLE Genomic Sequence for Oryza sativa, Nipponbare strain, clone
JOURNAL OSJNBa0067N01, from chromosome 3, complete sequence
REFERENCE 2 (bases 1 to 159636)
AUTHORS McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Center, Cold
Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
NY 11724, USA
REFERENCE 3 (bases 1 to 159636)
AUTHORS McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2001) Lita Annenberg Hazen Genome Center, Cold
Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
NY 11724, USA
REFERENCE 4 (bases 1 to 159636)
AUTHORS Palmer,L.E., de la Bastide,M., Spiegel,L., Preston,R., Kirchoff,K.,
King,L., Nascimento,L., Baker,J., Vil,M.D., Zutavern,T., Santos,L.,
Miller,B., Kuit,K., Cunnius,D.M., Rodriguez,S., Balija,V., Shah,R.,
Bahret,A., Yang,C., Bell,M., O'Shaughnessy,A., Dedhia,N. and
McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2001) Lita Annenberg Hazen Genome Center, Cold
Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY
11724, USA
REMARK Genomic Sequence for Oryza sativa, Nipponbare strain, clone
OSJNBa0067N01, from chromosome 3, complete sequence
COMMENT On Jun 20, 2001 this sequence version replaced gi:14209722.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.
FEATURES
source
1. 159636
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0067N01"
/clone_lib="HindIII"
/complement(851..931)
/note="transposon02 TNP2-like transposon protein gi|4"
1532..2626
/gene="OSJNBa0067N01.1"
/note="Hypothetical protein"
1532..2626
/gene="OSJNBa0067N01.1"
/note="similar to pir||T46123 hypothetical protein
T2J13.90 - Arabidopsis thaliana identities = 119/339
Evalue= 3e-44"
/codon_start=1
/product="Hypothetical protein"
/protein_id="AAK98722.1"
repeat_region
repeat_region
repeat_region
gene
1532..2626
/gene="OSJNBa0067N01.1"
/note="Putative retroelement"
Join(19417..20491,20573..20916)
/gene="OSJNBa0067N01.4"
/note="Similar to gb|AAK38301.1|AC079029.2 (AC079029)
putative reverse transcriptase [Oryza sativa] Identities =
170/488 Evalue= 3e-59"
/codon_start=1
/product="Putative retroelement"
/protein_id="AAK98725.1"
/db_xref="GI:15451601"

```

```

/db_xref="GI:15451598"
/translation="MIDAVHPFKVAVPSSPFFREBYTSAFRTESVNDVFWVLDTIT
LAHGAALVFRHGGGCGCNSARLFSYRLFADHLLPEQORAVAAALASPRGRLRPDR
GLLAAYAEATNSAFLCSLLKLDIEHLRYRPLKTRKLKASLVGSLGSAALG
QPFTALAAOQRLREVQAGDILRLGADGRKARHRSVARLRALVSFVTAALV
VAVGACIGVHILAAFAAPPMWGSFGRARRALVQLEAAAKGYIILNRDM
ETISRLVARVRDEGHMVALLRCLVEHRPAAGAGKGRVLQVLRQLSKNEESFRQL
DELEHFLCFTWTKARIMYMFMAAAR"
complement(3166..5556)
/gene="OSJNBa0067N01.2"
/note="Putative signal peptidase"
complement(join(3166..3273,4313..4387,5109..5274,
5366..5452,5534..5556))
/gene="OSJNBa0067N01.2"
/note="Similar to emb|CAC09470.1|(AL442113) putative
signal peptidase [Oryza sativa] Identities = 68/138
Evalue= 7e-33"
/codon_start=1
/product="Putative signal peptidase"
/protein_id="AAK98723.1"
/db_xref="GI:15451599"
/translation="MENGNLCELELGAVLAERSCLQYQFSGDVLVFKPCSDHRE
LFRKLIALPGEWQMLPGTPIIKIPGHCWVEGDNAACSDRSRSGFEVDGDKDSMG
GVRSSASGMIGPPRIPLIGLIGRVAHVIPPSKIGRVDTKPENRISPL"
7459..10855
/gene="OSJNBa0067N01.3"
/note="Unknown protein with similarity to
methyltransferases and transcriptional regulators"
Join(7459..9132,9247..9330,9641..9745,9839..9947,
10446..10597,10718..10855)
/gene="OSJNBa0067N01.3"
/note="Similar to dbj|BAA8534.1|(AP000969) Similar to
Arabidopsis thaliana DNA chromosome 4, BAC clone T5L19
(AL049481) [Oryza sativa] Identities = 571/759 Evalue=
0.0; Similar to sp|P25583|KAR4 YEAST KARYOGAMY PROTEIN
KAR4 Identities = 98/328 Evalue= 2e-27"
/codon_start=1
/product="Unknown protein with similarity to
methyltransferases and transcriptional regulators"
/protein_id="AAK98724.1"
/db_xref="GI:15451600"
/translation="MDSRPKIPRRSPDSDKDDNRNRGRDEKNDSSRTYGETD
CKERCNDKRGKSGAMGSDVDGDSRSDRSVDRSHETVHFVNDKQKAVEIKNLIHGVGQ
SDYQKQGLDNRNRNGTVKSRDAHDDKLGSRDNRNWTGKQEPGSGVDYLRSC
SQDSKASDSEKNAQERQDGGFGHVRDPRGSESTRGSTGYSDYDTSIE
IRPNNSLDFREGSVGRYDVGVAHRDVTYTGNGKVTNPEPDQSGSASMSQFPQH
CHKGRPRGGRGPRNGRSORGVNLTPIIMPFPGLCPGPMQHIQGNIPHSQPH
LLPGVFVPFGGPLLWPGAGVDNMLSVPNLPIPPVAGEHSFTGMCAGPNHIL
NQFGSIGAPTNMGLSFHQIGTQSRMWHCKPVGSGWTNRNRSPTKPSRGEON
DYSQNFVDTGRPNQFIRELDTISVADYFKRELIIQKDEIVAKSASPMMYKCDLR
QVLSPFFGTFKFDVLVDPPWRYVHRAPGITDHIYXWNGEEMNLIKIEAIDTSPF
VFLVGDGGLGEGQRCQKLFRCEDVCVNTKKSATPSLRHDSHTILOHKEHC
LMGIGTVRSTGDGHVIANIDTIIIADEPTDGTSTKPEDMYRIIEHFLGKRLEL
FGEDHNIRPGWLTILGKLSYFNKNKAYVKNFADKDGKVGQGRNPPEAPHLVVTP
RIEGLRPKSP9HK"
14781..14815
/note="misc_rsp03 similar to Rat tandem repeat mRNA"
complement(16651..16787)
/note="Galign_Ol2 MITE element from gb:U72728 Oryza 1"
complement(17170..17328)
/note="Wanderer_Os5 MITE element from gb:X13679 Oryza"
19417..20916
/gene="OSJNBa0067N01.4"
/note="Putative retroelement"
Join(19417..20491,20573..20916)
/gene="OSJNBa0067N01.4"
/note="Similar to gb|AAK38301.1|AC079029.2 (AC079029)
putative reverse transcriptase [Oryza sativa] Identities =
170/488 Evalue= 3e-59"
/codon_start=1
/product="Putative retroelement"
/protein_id="AAK98725.1"
/db_xref="GI:15451601"

```



```

QY 473 AAAGTTGGTCGGTTTCGTAATAAAAAAATTTAAAAATATATATTTTAAAAAACCCGAC 532
Db 1849 AGTAGTACCGCATCTGGTCGTATGTATA--TCAATAATCAATTTTATTATTTCGCC 1792
QY 533 CAACTTTAGTCGGTTTTTTGGTCGATTTTGGACGACCAAAAGTTGGTCGGTCGACCTTG 592
Db 1791 AATTAGCGACCAACTTTCGTCGCTTTTCTTGATAAACCAACTTTGGTCGCTAAAGTTTG 1732
QY 593 GTCGGTTTTTGGCGAATTTCTAGTACGTGACCGAAC 628
Db 1731 GTCGTTTTTTCGAAATTTCTAGTACGTGACCTACC 1696

RESULT 13
AP006654 120480 bp DNA linear PLN 03-FEB-2004
LOCUS Lotu corniculatus var. japonicus genomic DNA, chromosome 3,
DEFINITION clone:LjT31E21, TM0340, complete sequence.
ACCESSION AP006654
VERSION AP006654.1 GI:41688339
KEYWORDS HTG.
SOURCE Lotu corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotu corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
REFERENCE 1
AUTHORS Kato, T., Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. and Tabata, S.
TITLE Structural Analysis of a Lotus japonicus Genome. V. Sequence
Features and Mapping of Sixty-four TAC Clones Which Cover the 6.4
Mb Regions of the Genome
JOURNAL DNA Res. 10, 277-285 (2003)
REFERENCE 2 (bases 1 to 120480)
AUTHORS Sato, S.
TITLE Direct Submission
JOURNAL Submitted (12-NOV-2003) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337),
Fax: 81-438-52-3934)
FEATURES
    source
        1..120480
            /organism="Lotus corniculatus var. japonicus"
            /mol_type="genomic DNA"
            /variety="japonicus"
            /db_xref="taxon:34305"
            /chromosome="3"
            /clone="LjT31E21"
            /clone_lib="LjT library"
            /notes="TAC clone: TM0340, synonym: Lotus japonicus"

ORIGIN
Query Match 7.7%; Score 69.2; DB 8; Length 120480;
Best Local Similarity 65.6%; Pred. No. 0.0033;
Matches 118; Conservative 0; Mismatches 58; Indels 4; Gaps 1;

QY 36 TTAATTCAAAAAATTTGTAATAAAAAACCGACCAAAAGTTGATCGGTATTTAATTA 95
Db 17320 TAAGACTAAAAAATTTGCAAAAACAGGCAAGCAAAATAGCCTTGTGGGAAATG 17379
QY 96 T---GTAATAAAAGATTCTACTATCTGGGAATCGAACCGGGGTCTGTACTATGGCAAGA 151
Db 17380 TAAGACTAAAAAATTTGACACGACCGGGGATCGAACCGGGTCTGTACCGTGGCAGGG 17439
QY 152 TACTATTCTACCATAGACCAATGGTTCATTTTCTTTTATAGACTGTCTTTTATTGATTT 211
Db 17440 TACTATTCTACCATAGACCACTGGTGTCTCATGTTTCATTTACGCAATGCACTCTTTT 17499

RESULT 14
LLVTRNAG/c 7869 bp DNA linear PLN 12-SEP-1997
LOCUS

```

```

DEFINITION L. luteus tRNA-Gly genes and tRNA-Gly pseudogenes.
ACCESSION Z49255
VERSION Z49255.1 GI:2073451
KEYWORDS pseudogene; transfer RNA-Gly; tRNA-Gly gene.
SOURCE Lupinus luteus (yellow lupine)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae;
Lupinus.
REFERENCE 1
AUTHORS Nuc, P., Nuc, K., Szwejkowska-Kulinska, Z. and Pawelkiewicz, J.
TITLE Nucleotide sequence of nuclear tRNA-Gly genes and tRNA-Gly
pseudogenes from yellow lupin (Lupinus luteus): expression of the
tRNA(Gly) genes in vitro and in vivo
JOURNAL Acta Biochim. Pol. 44 (2), 259-274 (1997)
MEDLINE 98025190
PUBMED 9360715
REFERENCE 2 (bases 1 to 7869)
AUTHORS Nuc, P.W.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1995) Nuc P. W., University of Agriculture,
Biochemistry & Biotechnology, ul. Wolynska 35, Poznan, Poland, 60
637
COMMENT On May 8, 1997 this sequence version replaced gi:1924956.
FEATURES
    source
        1..7869
            /organism="Lupinus luteus"
            /mol_type="genomic DNA"
            /cultivar="ventus, line no. 099067, Poland, Wiatrowo,
            W.Swiecicki"
            /db_xref="taxon:3873"
            /clone="lambda G4"
            /clone_lib="EMBL3 genomic library"
            /cissue_type="epicotyl"
            /dev_stage="seedling"
            /rpt_type="TANDEM"
            /rpt_type="TANDEM"
            /gene="tRNA-Gly"
            /gene="tRNA-Gly"
            /gene="tRNA-Gly"
            /product="tRNA-Gly"
            /note="Codon recognized: GGC"
            /anticodon="(pos:364..366,aa:Gly)"
            /evidence="experimental"
            1075..1145
            /gene="tRNA-Gly"
            /gene="tRNA-Gly"
            /product="tRNA-Gly"
            /pseudo
            2897..2967
            /gene="tRNA-Gly"
            2897..2967
            /gene="tRNA-Gly"
            /product="tRNA-Gly"
            /evidence="experimental"
            3639..3709
            /gene="tRNA-Gly"
            3639..3709
            /gene="tRNA-Gly"
            /product="tRNA-Gly"
            /pseudo
            5461..5531
            /gene="tRNA-Gly"
            5461..5531
            /gene="tRNA-Gly"
            /product="tRNA-Gly"
            /evidence="experimental"
            6201..6271
            /gene="tRNA-Gly"

```

```
tRNA      6201.. 6271
/gene="tRNA-Gly"
/product="tRNA-Gly"
/pseudo

ORIGIN
Query Match      7.6%; Score 68.8; DB 8; Length 7869;
Best Local Similarity 59.0%; Pred. No. 0.0057;
Matches 118; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 36 TTAATTCAAAATAAATTTGTAATAAAACCAACCGAAGTTGATCGGTATTTAATTA 95
DB 5604 TTCATCTTAGAGATGTTAGTACCAACCAACACATATATGCTAATATCATCATAG 5545
QY 96 TGTATATAAAGATTCACTACTCTGGGAATCGAACCGGGTCTGTACTATGCAAGATACT 155
DB 5544 CTCAAAGTATATATGACACCGCGGGAATCGAACCGGGTCTGTACTATGCAAGATACT 5485
QY 156 ATTCTACCACTAGACCAATGCTTTCATTTGTTTAAAGACTGTCCTTTTATTGATTATAC 215
DB 5484 ATTCTACCACTAGACCACTGGTCTGTATGCTATTCTGCTTAGGTTTATCTTATTCAAC 5425
QY 216 TCTTTAATTATTTTTCGA 235
DB 5424 TAAATGCTTATATTCTACA 5405

RESULT 15
NAU40606/c
LOCUS      NAU40606      528 bp      DNA      linear      PLN 23-AUG-2001
DEFINITION Nicotiana alata isolate S2S2 S-like RNase gene, intron sequence.
ACCESSION  U40606
VERSION     U40606.1  GI:1272238
KEYWORDS
SOURCE      Nicotiana alata (Persian tobacco)
ORGANISM    Nicotiana alata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1 (bases 1 to 628)
            Dodds,P.N., Clarke,A.E. and Newbigin,E.
            Molecular characterisation of an S-like RNase of Nicotiana alata
            that is induced by phosphate starvation
            Plant Mol. Biol. 31 (2), 227-238 (1996)
JOURNAL     96343927
MEDLINE     8756589
REFERENCE   2 (bases 1 to 628)
            Dodds,P.N.
            Direct Submission
            Submitted (14-NOV-1995) Peter N Dodds, Botany, University of
            Melbourne, Parkville, Victoria, 3052, Australia
FEATURES
            source
            1..628
            Location/Qualifiers
            /organism="Nicotiana alata"
            /mol_type="genomic DNA"
            /isolate="S2S2"
            /db_xref="taxon:4087"
            1..628
            /gene="S-like RNase gene"
            1..628
            /gene="S-like RNase gene"
            275..481
            /gene="S-like RNase gene"
            /note="insertion sequence found in the intron region
            represents an allelic variant"

gene
intron
misc_feature

ORIGIN
Query Match      7.5%; Score 68; DB 8; Length 628;
Best Local Similarity 70.1%; Pred. No. 0.011;
Matches 103; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 600 TTTCGCCGAATTTCTAGTAGCAGCAACCTGTAAAGCTTCGGAGAAAATTTGTATATCT 659
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 06:24:54 ; Search time 444.342 Seconds
(without alignments)
10644.348 Million cell updates/sec

Title: US-10-009-966c-1_COPY_2255_3155
Perfect score: 901
Sequence: 1 tggctggaaattaccgaccca.....atgacttcattcgtattctc 901

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	899	99.8	3294	4 AAF25735	Aaf25735 Tobacco i
2	899	99.8	4135	4 AAF25742	Aaf25742 Tobacco p
3	899	99.8	4312	4 AAF25736	Aaf25736 Tobacco i
C 4	72.4	8.0	3294	4 AAF25735	Aaf25735 Tobacco i
C 5	72.4	8.0	4135	4 AAF25742	Aaf25742 Tobacco p
C 6	72.4	8.0	4312	4 AAF25736	Aaf25736 Tobacco i
C 7	64.8	7.2	1140	10 ADB87452	Adb87452 Transgene
C 8	64.8	7.2	1140	10 ADB87474	Adb87474 Transgene
C 9	64.8	7.2	1140	12 ADJ35261	Adj35261 Thale cre
C 10	64.8	7.2	1140	12 ADJ35283	Adj35283 Thale cre
C 11	64.8	7.2	1140	12 ADL97593	Adl97593 Stabilizin
C 12	64.8	7.2	1140	12 ADL96879	Adl96879 Stabilizin
C 13	64	7.1	5945	6 ABL32085	Ab132085 Human imm
C 14	62.8	7.0	15518	6 ABL34172	Ab134172 Human imm
C 15	62.8	7.0	15518	6 ABL34624	Ab134624 Human met
C 16	62.8	7.0	15518	6 ABL70607	Ab170607 Chemical
C 17	61.8	6.9	91608	10 AAL54538	Aal54538 Arabidops
C 18	57	6.3	6816	12 ADQ24856	Adq24856 Human sof
C 19	55.4	6.1	6239	6 ABL33776	Ab133776 Human imm
C 20	55.4	6.1	6239	6 ABK28297	Abk28297 DNA trans
C 21	55.2	6.1	5407	6 ABL34091	Ab134091 Human imm

22	54.8	6.1	8056	8 ABZ10246	Abz10246 Haematopo
C 23	54.6	6.1	368	4 AAS60046	Aas60046 Human can
24	53.4	5.9	8056	8 ABZ10100	Abz10100 Haematopo
25	53	5.9	6092	6 AAS61080	Aas61080 Human gen
C 26	52.8	5.9	8979	6 ABL32784	Ab132784 Human imm
C 27	52.8	5.9	8979	6 ABK31270	Abk31270 Signal tr
C 28	52.8	5.9	8979	6 ABL70231	Ab170231 Chemical
C 29	52.8	5.9	8979	6 ABL61178	Ab61178 Human gen
C 30	52.4	5.8	419	8 ABX46069	Abx46069 Bovine ES
C 31	52	5.7	50000	6 ABL55643	Ab155643 AnEPV gen
C 32	51.4	5.7	2000	8 ADA17938	Ada17938 Rice gene
C 33	51.2	5.7	358	4 AAI83451	Aai83451 Human pol
C 34	51	5.7	16914	6 ABL70316	Ab170316 Chemical
C 35	51	5.7	16914	6 AAS61254	Aas61254 Human gen
C 36	50.4	5.6	5756	6 ABL32586	Ab132586 Human imm
C 37	50.4	5.6	7234	6 ABK31228	Abk31228 Signal tr
38	50.4	5.6	15732	4 AAS45389	Aas45389 Chemical
39	50.4	5.6	15732	6 ABK28234	Abk28234 DNA trans
40	50.2	5.6	500	5 ADI73089	Adi73089 Human ova
41	50.2	5.6	500	5 ADL38222	Adl38222 Human ova
C 42	50	5.5	5997	6 ABL33625	Ab133625 Human imm
C 43	49.6	5.5	417	5 ABV09124	Abv09124 Human pro
C 44	49.6	5.5	1501	8 ABZ10042	Abz10042 Haematopo
C 45	49.6	5.5	1501	10 ADE84086	Ade84086 Human lym

ALIGNMENTS

RESULT 1
AAF25735
ID AAF25735 standard; DNA; 3294 BP.
AC AAF25735;
XX AAF25735;
DT 06-APR-2001 (first entry)
DE Tobacco invertase promoter SEQ ID NO 1.
KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
KW male-sterile plant; in vitro embryogenesis; seedless fruit;
KW transgenic plant; ds.
XX Nicotiana tabacum.
XX WO200077187-A2.
PD 21-DEC-2000.
XX 13-JUN-2000; 2000WO-DE001944.
XX 12-JUN-1999; 99DE-02009998.
PR 04-APR-2000; 2000DE-02005992.
PR 26-APR-2000; 2000DE-02007494.
PA (ROIT/) ROITSCH T.
PI Roitsch T;
XX WPI; 2001-080685/09.
DR New tapetum- and pollen-specific promoter from tobacco, useful for
XX preparing male sterile plants, particularly those with seedless fruits.
PS Claim 3; Page 64-65; 74pp; German.
CC This invention describes a novel nucleic acid (I) comprising a promoter
CC that is specific for tapetum and pollen. The invention also describes (a)
CC expression system containing at least one (i); (b) nucleic acid construct
CC (ii) containing (i) plus at least part of an expressible nucleic acid
CC (iii); (c) vector containing (i), the system of (a), or (ii); (d) cells,
CC particularly plant cells, containing (i), the system of (a), (ii) or the
CC vector of (c); (e) plants containing cells of (d); (f) seeds from the

parent
PCT

CC (I) can also be used to prepare transgenic plants that show increased or
 CC reduced production of endogenous materials, e.g. of plant hormones or
 CC proteins involved in provision of energy to developing tissue. (I)
 CC provide high level expression in a tissue- and time-specific manner, and
 CC do not require exogenous stimuli (e.g. temperature or chemicals) for
 CC regulation
 XX
 SQ Sequence 4312 BP; 1196 A; 827 C; 905 G; 1382 T; 0 U; 2 Other;

Query Match 99.8%; Score 899; DB 4; Length 4312;
 Best Local Similarity 99.8%; Pred. No. 1.1e-171;
 Matches 899; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGGTCGGAATACCGACCACTTTGGTCGTCGAATTAATTCAAAAAATTTGTAA 60
 DB 3273 TGGTCGGAATACCGACCACTTTGGTCGTCGAATTAATTCAAAAAATTTGTAA 3332
 QY 61 AAAAAACCGACCAAGTTGATCGGTATTTAATTAATTAATAAAGATTCACTATCTGG 120
 DB 3333 AAAAAACCGACCAAGTTGATCGGTATTTAATTAATTAATAAAGATTCACTATCTGG 3392
 QY 121 GAATCGAACCGGGTCTGTACTATGGAAGATATTTTACCACTAGACCATGGTTCA 180
 DB 3393 GAATCGAACCGGGTCTGTACTATGGAAGATATTTTACCACTAGACCATGGTTCA 3452
 QY 181 TTTTGTGTTTAAAGTCTGTTTATTTGATTTATATCTTTAATTAATTTTTCACGAA 240
 DB 3453 TTTTGTGTTTAAAGTCTGTTTATTTGATTTATATCTTTAATTAATTTTTCACGAA 3512
 QY 241 ATAACCGACCAAGTTGGTCGATTTTATTAAGTAAATTTACTTACCAAGTTGGTCG 300
 DB 3513 ATAACCGACCAAGTTGGTCGATTTTATTAAGTAAATTTACTTACCAAGTTGGTCG 3572
 QY 301 ATTTTATTAATGATCGCCGAATTAACCGCAATTTTCGTAGGTTTATTAATTA 360
 DB 3573 ATTTTATTAATGATCGCCGAATTAACCGCAATTTTCGTAGGTTTATTAATTA 3632
 QY 361 TTTTATTTATTTAAATGAAAACTAACCAAGTTAGTCGGTTTCTGAAAAATTAAT 420
 DB 3633 TTTTATTTATTTAAATGAAAACTAACCAAGTTAGTCGGTTTCTGAAAAATTAAT 3692
 QY 421 TCGCGGACTCAAAATAGTTTCCGCAATTTTCGCGCAAGAAACCGACCAAGTTGG 480
 DB 3693 TCGCGGACTCAAAATAGTTTCCGCAATTTTCGCGCAAGAAACCGACCAAGTTGG 3752
 QY 481 TCGGTTTCGTAATAAAAAAATTTTAAATATATATTTTAAACCAAGCACTTTA 540
 DB 3753 TCGGTTTCGTAATAAAAAAATTTTAAATATATATTTTAAACCAAGCACTTTA 3812
 QY 541 GTCGGTTTTCGTCGATTTTTCACCGCAAGTTGTCGTCGACCTTGGTCGGTTT 600
 DB 3813 GTCGGTTTTCGTCGATTTTTCACCGCAAGTTGTCGTCGACCTTGGTCGGTTT 3872
 QY 601 TTGCGGAATTTAGTAGTACCGACCTGTAGCTTCGGGAAATTTTGTATATGA 660
 DB 3873 TTGCGGAATTTAGTAGTACCGACCTGTAGCTTCGGGAAATTTTGTATATGA 3932
 QY 661 TATGTGTATATCCTTAAATGATTAATTTAAAGAACGNNCCACCTGAACTACTAGAAGCC 720
 DB 3933 TATGTGTATATCCTTAAATGATTAATTTAAAGAACGNNCCACCTGAACTACTAGAAGCC 3992
 QY 721 TTTAGGGCACTAGATGAGCAGAAATACGTGTTCTGTCGCTGAAAAATTTGATCGG 780
 DB 3993 TTTAGGGCACTAGATGAGCAGAAATACGTGTTCTGTCGCTGAAAAATTTGATCGG 4052
 QY 781 CCTATGATGTAAGTACTTCTGTCCTTAAATCAGAGTTTTCGACTTCGAGTCCAGATA 840
 DB 4053 CCTATGATGTAAGTACTTCTGTCCTTAAATCAGAGTTTTCGACTTCGAGTCCAGATA 4112
 QY 841 TAAACTATAGATCGCTTTTATAGCACTTTTAAAGACTATGATTCATCTGATTTCT 900
 DB 4113 TAAACTATAGATCGCTTTTATAGCACTTTTAAAGACTATGATTCATCTGATTTCT 4172

QY 901 C 901
 DB 4173 C 4173

RESULT 4
 AAF25735/c
 ID AAF25735 standard; DNA; 3294 BP.
 XX AC AAF25735;
 XX DT 06-APR-2001 (first entry)
 XX DE Tobacco invertase promoter SEQ ID NO 1.
 XX KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
 KW male-sterile plant; in vitro embryogenesis; seedless fruit;
 XX transgenic plant; ds.
 XX OS Nicotiana tabacum.
 XX FN WO200077187-A2.
 XX PD 21-DEC-2000.
 XX PF 13-JUN-2000; 2000WO-DE001944.
 XX PR 12-JUN-1999; 99DE-02009998.
 PR 04-APR-2000; 2000DE-02005992.
 PR 26-APR-2000; 2000DE-02007494.
 XX (ROIT/) ROITSCH T.
 XX PI Roitsch T;
 XX WPI; 2001-080685/09.
 XX New tapetum- and pollen-specific promoter from tobacco, useful for
 preparing male sterile plants, particularly those with seedless fruits.
 PS Claim 3; Page 64-65; 74pp; German.

This invention describes a novel nucleic acid (I) comprising a promoter
 that is specific for tapetum and pollen. The invention also describes (a)
 expression system containing at least one (I); (b) nucleic acid construct
 (II) containing (I) plus at least part of an expressible nucleic acid
 (III); (c) vector containing (II), the system of (a), or (II); (d) cells,
 particularly plant cells, containing (II), the system of (a), (II) or the
 vector of (c); (e) plants containing cells of (d); (f) seeds from the
 plants of (e); (g) hybrid seed produced by crossing a male sterile plant
 of (e); (h) production of male sterile plants by introducing (II) into a
 cell and regeneration to a plant; (i) restorer plants containing in one
 (or preferably many) cells a construct (IIa) of (I) and a sequence that
 encodes an invertase different from the plants endogenous invertase; (j)
 plant that contains in one (or preferably many) cells both (II) and (IIa)
 ; (k) seeds from plants of (i) and (j); (l) fruits, particularly
 seedless, produced by plants of (e), (i) and (j); and (m) method for
 cloning a promoter that is functionally homologous with (I). Constructs
 containing (I) and an invertase-encoding sequence are used to produce
 male-sterile plants (by co-suppression or antisense techniques) for
 preparation of hybrids, while constructs that contain (I) and a
 heterologous invertase-encoding sequence are used to prepare restorer
 plants (which allow propagation of the male-sterile plants). Seeds from
 these plants are used for in vitro embryogenesis of haploid or (double)
 diploid plants, and the new plants particularly produce seedless fruits.
 (I) can also be used to prepare transgenic plants that show increased or
 reduced production of endogenous materials, e.g. of plant hormones or
 proteins involved in provision of energy to developing tissue. (I)
 provide high level expression in a tissue- and time-specific manner, and
 do not require exogenous stimuli (e.g. temperature or chemicals) for
 regulation

Sequence 3294 BP; 983 A; 597 C; 644 G; 1064 T; 0 U; 6 Other;
 SQ

```

Query Match      8.0%; Score 72.4; DB 4; Length 3294;
Best Local Similarity 60.3%; Pred. No. 3.2e-05;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCGCGGAATTAAACGACCAATTTTGGTAGGTTTTTTTAATATTATTTTA- 366
DB 2819 TCAAAAATCGACCAAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760

QY 367 ---TTTATTTTAAATGAAAACCTAACCAAGTTAGTCGGTTCTTGAAACATAAATTCG 423
DB 2759 ATTTTATTTTATTTTACGAAACCGACCAACTTTGGTCGGTTTTCTTTGGCGCAAAATGC 2700

QY 424 CGGGAATCAAAAATAGTTTCCCGCATTTTGGCCCAAGAAAACCGACCAAAAGTTGGTCG 483
DB 2699 GGGAACTATTTTGTAGTCGCGGAATTTATGTTTCAAGAAACCGACTAACTTTGGTTA 2640

QY 484 GTTTCGTAAAAAATAAATTTTAAATAATATATTTTAAATAACCGACCAACTTTAGTC 543
DB 2639 GTTTTCAATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2584

QY 544 GGTTTTTTGGTCGATTTTTCGA 565
DB 2583 GGTAAATTCGGCGGATCAATTA 2562

RESULT 5
AAF25742/c
ID AAF25742 standard; DNA; 4135 BP.
XX
AC AAF25742;
XX
DT 06-APR-2001 (first entry)
XX
DE Tobacco promoter/invertase NIN88 DNA fragment SEQ ID NO 8.
XX
KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
KW male-sterile plant; in vitro embryogenesis; seedless fruit;
KW transgenic plant; ds.
XX
OS Nicotiana tabacum.
XX
PN WO200077187-A2.
XX
PD 21-DEC-2000.
XX
PF 13-JUN-2000; 2000WO-DE001944.
XX
PR 12-JUN-1999; 99DE-02009998.
PR 04-APR-2000; 2000DE-02005992.
PR 26-APR-2000; 2000DE-02007494.
XX
(ROIT/) ROITSCH T.
XX
PI Roitsch T;
XX
DR WPI; 2001-080685/09.
XX
PT New tapetum- and pollen-specific promoter from tobacco, useful for
PT preparing male sterile plants, particularly those with seedless fruits.
XX
PS Example 4.3; Page 68-70; 74pp; German.
XX
CC This invention describes a novel nucleic acid (I) comprising a promoter
CC that is specific for tapetum and pollen. The invention also describes (a)
CC expression system containing at least one (i); (b) nucleic acid construct
CC (II) containing (I) plus at least part of an expressible nucleic acid
CC (III); (c) vector containing (II), the system of (a), or (II); (d) cells,
CC particularly plant cells, containing (I), the system of (a), (II) or the
CC vector of (c); (e) plants containing cells of (d); (f) seeds from the
CC plants of (e); (g) hybrid seed produced by crossing a male sterile plant
CC of (e); (h) production of male sterile plants by introducing (II) into a
CC cell and regeneration to a plant; (i) restorer plants containing in one

```

```

(or preferably many) cells a construct (IIa) of (I) and a sequence that
encodes an invertase different from the plants endogenous invertase; (j)
plant that contains in one (or preferably many) cells both (II) and (IIa)
; (k) seeds from plants of (i) and (j); (l) fruits, particularly
; seedless, produced by plants of (e), (i) and (j); and (m) method for
CC cloning a promoter that is functionally homologous with (I). Constructs
CC containing (I) and an invertase-encoding sequence are used to produce
CC male-sterile plants (by co-suppression or antisense techniques) for
CC preparation of hybrids, while constructs that contain (I) and a
CC heterologous invertase-encoding sequence are used to prepare restorer
CC plants (which allow propagation of the male-sterile plants). Seeds from
CC these plants are used for in vitro embryogenesis of haploid or (double)
CC diploid plants, and the new plants particularly produce seedless fruits.
CC (I) can also be used to prepare transgenic plants that show increased or
CC reduced production of endogenous materials, e.g. of plant hormones or
CC proteins involved in provision of energy to developing tissue. (I)
CC provide high level expression in a tissue- and time-specific manner, and
CC do not require exogenous stimuli (e.g. temperature or chemicals) for
CC regulation
XX
SQ Sequence 4135 BP; 1212 A; 774 C; 825 G; 1318 T; 0 U; 6 Other;

Query Match      8.0%; Score 72.4; DB 4; Length 4135;
Best Local Similarity 60.3%; Pred. No. 3.3e-05;
Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;

QY 308 TAAATGATCGCGGAATTAAACGACCAATTTTGGTAGGTTTTTTTAATATTATTTTA- 366
DB 2819 TCAAAAATCGACCAAAAACCGACTAAAGTTGGTCGGTTTTTAAATATATTTTAA 2760

QY 367 ---TTTATTTTAAATGAAAACCTAACCAAGTTAGTCGGTTCTTGAAACATAAATTCG 423
DB 2759 ATTTTATTTTATTTTACGAAACCGACCAACTTTGGTCGGTTTTCTTTGGCGCAAAATGC 2700

QY 424 CGGGAATCAAAAATAGTTTCCCGCATTTTGGCCCAAGAAAACCGACCAAAAGTTGGTCG 483
DB 2699 GGGAACTATTTTGTAGTCGCGGAATTTATGTTTCAAGAAACCGACTAACTTTGGTTA 2640

QY 484 GTTTCGTAAAAAATAAATTTTAAATAATATATTTTAAATAACCGACCAACTTTAGTC 543
DB 2639 GTTTTCAATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2584

QY 544 GGTTTTTTGGTCGATTTTTCGA 565
DB 2583 GGTAAATTCGGCGGATCAATTA 2562

RESULT 6
AAF25736/c
ID AAF25736 standard; DNA; 4312 BP.
XX
AC AAF25736;
XX
DT 06-APR-2001 (first entry)
XX
DE Tobacco invertase promoter SEQ ID NO 2.
XX
KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
KW male-sterile plant; in vitro embryogenesis; seedless fruit;
KW transgenic plant; ds.
XX
OS Nicotiana tabacum.
XX
PN WO200077187-A2.
XX
PD 21-DEC-2000.
XX
PF 13-JUN-2000; 2000WO-DE001944.
XX
PR 12-JUN-1999; 99DE-02009998.
PR 04-APR-2000; 2000DE-02005992.
PR 26-APR-2000; 2000DE-02007494.
XX

```

PA (ROIT/) ROITSCH T.
 XX
 XX
 PI Roitsch T;
 XX
 XX
 DR WPI; 2001-080685/09.
 XX
 XX New tapetum- and pollen-specific promoter from tobacco, useful for
 PT preparing male sterile plants, particularly those with seedless fruits.
 XX
 XX
 PS Claim 6; Page 65-66; 74pp; German.
 XX
 XX This invention describes a novel nucleic acid (I) comprising a promoter
 CC that is specific for tapetum and pollen. The invention also describes (a)
 CC expression system containing at least one (I); (b) nucleic acid construct
 CC (II) containing (I) plus at least part of an expressible nucleic acid
 CC (III); (c) vector containing (II), the system of (a), or (II); (d) cells,
 CC particularly plant cells, containing (I), the system of (a), (II) or the
 CC vector of (c); (e) plants containing cells of (d); (f) seeds from the
 CC plants of (e); (g) hybrid seed produced by crossing a male sterile plant
 CC of (e); (h) production of male sterile plants by introducing (II) into a
 CC cell and regeneration to a plant; (i) restorer plants containing in one
 CC (or preferably many) cells a construct (IIa) of (I) and a sequence that
 CC encodes an invertase different from the plants endogenous invertase; (j)
 CC plant that contains in one (or preferably many) cells both (II) and (IIa)
 CC ; (k) seeds from plants of (i) and (j); (l) fruits, particularly
 CC seedless, produced by plants of (e), (i) and (j); and (m) method for
 CC cloning a promoter that is functionally homologous with (I). Constructs
 CC containing (I) and an invertase-encoding sequence are used to produce
 CC male-sterile plants (by co-suppression or antisense techniques) for
 CC preparation of hybrids, while constructs that contain (I) and a
 CC heterologous invertase-encoding sequence are used to prepare restorer
 CC plants (which allow propagation of the male-sterile plants). Seeds from
 CC these plants are used for in vitro embryogenesis of haploid or (double)
 CC diploid plants, and the new plants particularly produce seedless fruits.
 CC (I) can also be used to prepare transgenic plants that show increased or
 CC reduced production of endogenous materials, e.g. of plant hormones or
 CC proteins involved in provision of energy to developing tissue. (I)
 CC provide high level expression in a tissue- and time-specific manner, and
 CC do not require exogenous stimuli (e.g. temperature or chemicals) for
 CC regulation
 XX
 XX
 SQ Sequence 4312 BP; 1196 A; 827 C; 905 G; 1382 T; 0 U; 2 Other;
 Query Match 8.0%; Score 72.4; DB 4; Length 4312;
 Best Local Similarity 60.3%; Pred. No. 3.3e-05;
 Matches 158; Conservative 0; Mismatches 96; Indels 8; Gaps 2;
 QY 308 TAAATGATCCGCGAATTAAACCGACCAATTTTGGTAGGTTTTTTTAAATATTAATTTTAA- 366
 DB 3837 TCARAAATCGACCAAAACCGACTAAAGTTGGTGGTTTTTAAATATATTTTAA 3778
 QY 367 ---TTTATTTAATGAAACTAACCAAGTAGTCGGTTCTTGAAACATAAATTCG 423
 DB 3777 ATTTTTTTTTTTTTTACGAAACCGACCAACTTTGGTGGTTTTTTTGGCGCAAAATGC 3718
 QY 424 CGGCACTCAAAATAGTTTCGGCATTTTTCGCGCAAGAAACCGACCAAGTTGGTGC 483
 DB 3717 GGGAACTATTTTGGTCCGCGAAATTTATGTTCAAGAACCGACTAATTTGGTTA 3658
 QY 484 GTTTCGTAATAAAAAAATTTTAAATAATATTTTAAATAAACCGACCAATTTAGTC 543
 DB 3657 GTTTTTCAATTAATAA---TAAATAATTAATTAATAAATAACCTACCAAAATTTGGTC 3602
 QY 544 GGTTTTTTGGTCGATTTTGA 565
 DB 3601 GGTTAATTCGGCGGATCATTTA 3580
 RESULT 7
 ADB87452/c
 ID ADB87452 standard; DNA; 1140 BP.
 XX
 XX
 AC ADB87452;
 04-DEC-2003 (first entry)
 Transgene expression regulatory element, STAR A13.
 gene transcription; regulatory; variety fragment; STAR element;
 transgene expression; ds.
 Unidentified.
 WO2003004704-A2.
 16-JAN-2003.
 14-JUN-2002; 2002WO-NL000390.
 04-JUL-2001; 2001EP-00202581.
 05-JUL-2001; 2001US-0303199F.
 (CHRO-) CHROMAGENICS BV.
 Otte AP, Kruckeberg AL;
 WPI; 2003-229412/22.
 Selecting a DNA sequence with a gene transcription modulating quality by
 providing a transcription system with a variety fragment-comprising
 PT vectors and performing a selection step in the transcription system.
 XX
 XX Claim 43; Fig 26; 216pp; English.
 PS The invention relates to DNA sequences with gene transcription regulatory
 CC qualities and methods for the detection and use of the regulatory DNA
 CC sequences. The invention further comprises providing a transcription
 CC system with a variety fragment-comprising vectors; and performing a
 CC selection step in the transcription system in order to identify a
 CC fragment comprising the DNA sequence with the gene transcription
 CC modulating activity. This polynucleotide represents a STAR element used
 CC in the method of the invention. The STAR element is shown to improve
 CC transgene expression.
 XX
 SQ Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
 Query Match 7.2%; Score 64.8; DB 10; Length 1140;
 Best Local Similarity 60.0%; Pred. No. 0.00096;
 Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 98 TAAATAAAAGATTTCATCTCTGGGAATCGAACCGGGGTCTGTACTATGGCAAGATATAT 157
 DB 366 TCATAGCTCAATGCACACGCGGGAATCGAACCGGGTCTGTACTATGGCAAGGTACTAT 307
 QY 158 TCTACCACTAGACCATGGTTCATTTGTTTAAAGACTGCTTTTATTGTTTATCTC 217
 DB 306 TCTACCACTAGACCACTGGTCTTTTGTATTAAACACGACTAAATATTAATATAGAT 247
 QY 218 TTTATATTATTTTTCACGAAATTAACCGACCAAAAGTTGGTTCGATTTTATTAATAAGTA 277
 DB 246 ATTATAAGTATCATGTATTAACCAACATCTTTTGTAGTTGTACCACTACGAAGTA 187
 RESULT 8
 ADB87474/c
 ID ADB87474 standard; DNA; 1140 BP.
 XX
 XX
 AC ADB87474;
 04-DEC-2003 (first entry)
 Transgene expression regulatory element, STAR A35.
 gene transcription; regulatory; variety fragment; STAR element;
 transgene expression; ds.
 XX

OS Unidentified.
 XX WO2003004704-A2.
 XX
 XX PD 16-JAN-2003.
 XX
 XX PF 14-JUN-2002; 2002WO-NL000390.
 XX
 XX PR 04-JUL-2001; 2001EP-0020581.
 XX PR 05-JUL-2001; 2001US-0303199P.
 XX
 XX PA (CHRO-) CHROMAGENICS BV.
 XX
 XX PI Otte AP, Kruckeberg AL;
 XX DR WPI; 2003-229412/22.
 XX
 XX Selecting a DNA sequence with a gene transcription modulating quality by
 PT providing a transcription system with a variety fragment-comprising
 PT vectors and performing a selection step in the transcription system.
 XX
 XX PS Claim 43; Fig 26; 216pp; English.
 XX
 XX The invention relates to DNA sequences with gene transcription regulatory
 CC qualities and methods for the detection and use of the regulatory DNA
 CC sequences. The invention further comprises providing a transcription
 CC system with a variety fragment-comprising vectors; and performing a
 CC selection step in the transcription system in order to identify a
 CC fragment comprising the DNA sequence with the gene transcription
 CC modulating activity. This polynucleotide represents a STAR element used
 CC in the method of the invention. The STAR element is shown to improve
 CC transgene expression.
 XX
 XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
 SQ
 Query Match 7.2%; Score 64.8; DB 10; Length 1140;
 Best Local Similarity 60.0%; Pred. No. 0.00096;
 Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 98 TAATAAAAGATTTCACCTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 157
 DB 366 TCATAGCTCAATGCACGCGCGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 307
 QY 158 TCTACCACTAGACCATGGTTCATTTTGTGTTTAAAGACTGCTTTTATTTGATTATATCTC 217
 DB 306 TCTACCACTAGACCATGGTTCATTTTGTGTTTAAAGACTGCTTTTATTTGATTATATAGAT 247
 QY 218 TTTAATTATATTTTGCACGAAATACCGCAAAAGTTGGTCGATTTTATTTAAAGTA 277
 DB 246 ATTATAAAGTATCATGTATTAACACACACAAATCTTTGTAGTTTGTACCACTACGAAGTA 187
 RESULT 9
 ADJ35261/c
 ID ADJ35261 standard; DNA; 1140 BP.
 XX
 XX AC ADJ35261;
 XX
 XX DT 22-APR-2004 (first entry)
 XX
 XX DE Thale cress stabilising anti-repression, STAR, element #13.
 XX
 XX KW STAR affiliated proteinaceous molecule; post translational modification;
 XX KW thale cress; stabilising anti-repression; STAR; STAR element; ds.
 XX
 XX OS Arabidopsis thaliana.
 XX
 XX PN WO2003106674-A2.
 XX
 XX PD 24-DEC-2003.
 XX
 XX PF 30-MAY-2003; 2003WO-NL000410.
 XX
 XX PR 14-JUN-2002; 2002EP-00077344.
 XX
 XX PA (CHRO-) CHROMAGENICS BV.
 XX
 XX PI Otte AP, Kruckeberg AL, Satijn DPE;
 XX DR WPI; 2004-082195/08.

PR 14-JUN-2002; 2002EP-00077344.
 XX
 XX PA (CHRO-) CHROMAGENICS BV.
 XX
 XX PI Otte AP, Kruckeberg AL, Satijn DPE;
 XX DR WPI; 2004-082195/08.
 XX
 XX Producing proteinaceous molecules in cells by selecting a cell, providing
 PT a nucleic acid encoding a proteinaceous molecule with an Stabilizing Anti
 PT -Repression sequence and expressing proteinaceous molecule.
 XX
 XX PS Example 13; SEQ ID NO 97; 177pp; English.
 XX
 XX The invention relates to a method of producing a proteinaceous molecule
 CC (I) in a cell comprising selecting a cell for its suitability for
 CC producing (I), providing a nucleic acid encoding (I) with a nucleic acid
 CC comprising a Stabilising Anti-Repression (STAR) sequence, expressing the
 CC resulting nucleic acid in the cell and collecting (I). The method is
 CC useful for producing (I). A cell line (II) provided with a nucleic acid
 CC comprising a STAR sequence is useful for producing (I). (II) Enables
 CC production of affiliated proteinaceous molecule, as cell carries out
 CC proper post-translational modifications of produced proteins. The present
 CC sequence represents a thale cress stabilising anti-repression, STAR,
 CC element.
 XX
 XX Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
 SQ
 Query Match 7.2%; Score 64.8; DB 12; Length 1140;
 Best Local Similarity 60.0%; Pred. No. 0.00096;
 Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 98 TAATAAAAGATTTCACCTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 157
 DB 366 TCATAGCTCAATGCACGCGCGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 307
 QY 158 TCTACCACTAGACCATGGTTCATTTTGTGTTTAAAGACTGCTTTTATTTGATTATATCTC 217
 DB 306 TCTACCACTAGACCATGGTTCATTTTGTGTTTAAAGACTGCTTTTATTTGATTATATAGAT 247
 QY 218 TTTAATTATATTTTGCACGAAATACCGCAAAAGTTGGTCGATTTTATTTAAAGTA 277
 DB 246 ATTATAAAGTATCATGTATTAACACACACAAATCTTTGTAGTTTGTACCACTACGAAGTA 187
 RESULT 10
 ADJ35283/c
 ID ADJ35283 standard; DNA; 1140 BP.
 XX
 XX AC ADJ35283;
 XX
 XX DT 22-APR-2004 (first entry)
 XX
 XX DE Thale cress stabilising anti-repression, STAR, element #35.
 XX
 XX KW STAR affiliated proteinaceous molecule; post translational modification;
 XX KW thale cress; stabilising anti-repression; STAR; STAR element; ds.
 XX
 XX OS Arabidopsis thaliana.
 XX
 XX PN WO2003106674-A2.
 XX
 XX PD 24-DEC-2003.
 XX
 XX PF 30-MAY-2003; 2003WO-NL000410.
 XX
 XX PR 14-JUN-2002; 2002EP-00077344.
 XX
 XX PA (CHRO-) CHROMAGENICS BV.
 XX
 XX PI Otte AP, Kruckeberg AL, Satijn DPE;
 XX DR WPI; 2004-082195/08.

XX Producing proteinaceous molecules in cells by selecting a cell, providing
PT a nucleic acid encoding a proteinaceous molecule with an Stabilizing Anti-
PT -Repressor sequence and expressing proteinaceous molecule.
XX
PS Example 13; SEQ ID NO 119; 177pp; English.
XX
CC The invention relates to a method of producing a proteinaceous molecule
CC (I) in a cell comprising selecting a cell for its suitability for
CC producing (I), providing a nucleic acid encoding (I) with a nucleic acid
CC comprising a Stabilizing Anti-Repressor (STAR) sequence, expressing the
CC resulting nucleic acid in the cell and collecting (I). The method is
CC useful for producing (I). A cell line (II) provided with a nucleic acid
CC comprising a STAR sequence is useful for producing (I). (II) Enables
CC production of affiliated proteinaceous molecule, as cell carries out
CC proper post-translational modifications of produced proteins. The present
CC sequence represents a thale cress stabilising anti-repression, STAR,
CC element.
XX
SQ Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
Query Match 7.2%; Score 64.8; DB 12; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 157
Db 366 TCATAGCTCAATGACCAACCGCGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 307
QY 158 TCTACCACTAGACCACTGGTCTCTTTTGTAAAGACTGCTTTTATTGATTATAT 217
Db 306 TCTACCACTAGACCACTGGTCTCTTTTGTAAAGACTGCTTTTATTGATTATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATAACCGACCAAGTTGTCGATTTTATAAAGTA 277
Db 246 ATTATAAGTATCATGTATAAACAACAACAATCTTTTGTAGTTTGTACCAACTACGAAGTA 187
RESULT 11
ADL97593/c
ID ADL97593 standard; DNA; 1140 BP.
XX AC ADL97593;
XX
DT 20-MAY-2004 (first entry)
XX
DE Stabilizing Anti-Repressor DNA sequence, A35.
XX
KW protein expression unit; Stabilizing Anti-Repressor; STAR; heterologous;
KW multimeric; transgene expression; ds.
XX
OS Arabidopsis sp.
XX
PN WO2003106684-A2.
XX
PD 24-DEC-2003.
XX
PF 13-JUN-2003; 2003WO-NL000432.
XX
PR 14-JUN-2002; 2002EP-00077350.
XX
PA (CHRO-) CHROMAGENICS BV.
XX
PI Otte AP, Kruckeberg AL, Sewalt RGAB;
XX
DR WPI; 2004-082197/08.
XX
CC Obtaining (M1) a cell which expresses two or more proteins, by providing
PT cell with protein expression units encoding proteins, comprising protein
PT expression units with Stabilizing Anti-Repressor sequence.
XX
PS Example 14; SEQ ID NO 119; 213pp; English.
XX

CC The invention relates to a novel method for obtaining a cell which
CC expresses two or more proteins. The method comprises providing the cell
CC with two or more protein expression units encoding the two or more
CC proteins, having at least two of the protein expression units with at
CC least one Stabilizing Anti-Repressor (STAR) sequence. The method provides
CC increased predictability in the production of recombinant cell lines that
CC effectively produce heterologous multimeric proteins of interest,
CC increase the yield of heterologous multimeric proteins, stable expression
CC of heterologous multimeric proteins and favorable transgene expression.
CC This polynucleotide sequence represents the DNA of a Stabilizing Anti-
XX Receptor (STAR) sequence of the invention.
XX
SQ Sequence 1140 BP; 357 A; 221 C; 186 G; 376 T; 0 U; 0 Other;
Query Match 7.2%; Score 64.8; DB 12; Length 1140;
Best Local Similarity 60.0%; Pred. No. 0.00096;
Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 98 TAATAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 157
Db 366 TCATAGCTCAATGACCAACCGCGGAATCGAACCGGGTCTGTACTATGGCAAGATACTAT 307
QY 158 TCTACCACTAGACCACTGGTCTCTTTTGTAAAGACTGCTTTTATTGATTATAT 217
Db 306 TCTACCACTAGACCACTGGTCTCTTTTGTAAAGACTGCTTTTATTGATTATATAGAT 247
QY 218 TTTAATTATATTTTGCACGAAATAACCGACCAAGTTGTCGATTTTATAAAGTA 277
Db 246 ATTATAAGTATCATGTATAAACAACAACAATCTTTTGTAGTTTGTACCAACTACGAAGTA 187
RESULT 12
ADL96879/c
ID ADL96879 standard; DNA; 1140 BP.
XX AC ADL96879;
XX
DT 20-MAY-2004 (first entry)
XX
DE Stabilizing Anti-Repressor DNA sequence, A13.
XX
KW protein expression unit; Stabilizing Anti-Repressor; STAR; heterologous;
KW multimeric; transgene expression; ds.
XX
OS Arabidopsis sp.
XX
PN WO2003106684-A2.
XX
PD 24-DEC-2003.
XX
PF 13-JUN-2003; 2003WO-NL000432.
XX
PR 14-JUN-2002; 2002EP-00077350.
XX
PA (CHRO-) CHROMAGENICS BV.
XX
PI Otte AP, Kruckeberg AL, Sewalt RGAB;
XX
DR WPI; 2004-082197/08.
XX
CC Obtaining (M1) a cell which expresses two or more proteins, by providing
PT cell with protein expression units encoding proteins, comprising protein
PT expression units with Stabilizing Anti-Repressor sequence.
XX
PS Example 14; SEQ ID NO 97; 213pp; English.
XX
CC The invention relates to a novel method for obtaining a cell which
CC expresses two or more proteins. The method comprises providing the cell
CC with two or more protein expression units encoding the two or more
CC proteins, having at least two of the protein expression units with at
CC least one Stabilizing Anti-Receptor (STAR) sequence. The method provides
CC increased predictability in the production of recombinant cell lines that
CC effectively produce heterologous multimeric proteins of interest,

PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIG-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 XX Claim 1; SEQ ID NO 2145; 32pp + Sequence Listing; German.
 PS The present invention provides a number of human immune system associated
 XX genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 XX
 XX Sequence 15518 BP; 3047 A; 362 C; 4183 G; 7926 T; 0 U; 0 Other;
 Query Match 7.0%; Score 62.8; DB 6; Length 15518;
 Best Local Similarity 46.6%; Pred. NO. 0.0032;
 Matches 234; Conservative 0; Mismatches 267; Indels 1; Gaps 1;
 QY 34 AATTAAATTCAAAAAATATTTGTAATAAAAAACCGACCAAGTTGATCGGTATTTAAT 93
 Db 942 AACTACGACCTAAATAAATAAAAAACAACTTCAAAATCTATTATAAAATTTCTTT 883
 QY 94 TATGTATAAAGATTCACATCTCTGGATCGAACCGGGGTCTGTACTATGCGCAAGATA 153
 Db 882 ACTAAAAAATAAAAAAATTTCTAAACAAATCAACACATATTAAATTTAATAAATA 823
 QY 154 CTATCTACCTAGACCATTTGTTTAAAGACTGTCTTTTAAAGACTGTCTTTTAAATTTAT 213
 Db 822 AAAAATACTATTAAAAATTAATACTAATACTCTCCATCTATTCTATATTTCTTAAAT 763
 QY 214 ACTCTTTAATTAATTTTGCACGAAATAAACCGACCAAGTTGGTTCGATTTTATAAAA 273
 Db 762 AATTATATCTTAA 704
 QY 274 AGTAAATTTACTTACCAAAATTAATACTTAAATTAATACTTAAATTAATACTAATA 333
 Db 703 AAAATCTTTCAAAAATAAATACTTAAATTAATACTTAAATTAATACTAATA 644
 QY 334 AATTTCGTAGTTTCTTTTAAATTAATACTTAAATTAATACTTAAATTAATACTAATA 584
 Db 643 CGGTAAATTTCTTTTAAATTAATACTTAAATTAATACTTAAATTAATACTAATA 524
 QY 394 GTTAGTCGGTTTCTTTGAAACATATAATTTTCGGGGACTCAAAATAGTTTCCGGATTTT 453
 Db 583 CTCGCTAGCTTACCAAAATTAATACTTAAATTAATACTTAAATTAATACTAATA 524
 QY 454 GCGCCAAAGAAACCGACCAAGTTGTCGGTTTCTGTAATAAAAAAATAAAAAA 513
 Db 523 AACAAATAAATCTATCAAAAAAATAAAAAAATAAAAAAATAAAAAA 464
 QY 514 TATATTTTAAACCGACCAA 535
 Db 463 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 464
 RESULT 15
 ABL34624/c
 ID ABL34624 standard; DNA; 15518 BP.
 XX
 XX ABL34624;
 XX
 DT 26-MAR-2002 (first entry)

XX Human metastasis associated gene SEQ ID NO: 177.
 DE
 XX Metastasis associated gene; cytostatic; gene therapy; cancer;
 KW cytosine methylation; gene; ds.
 XX Homo sapiens.
 OS
 XX WO200177376-A2.
 PN 18-OCT-2001.
 XX
 XX 06-APR-2001; 2001WO-EP003970.
 PF
 XX 06-APR-2000; 2000DE-01019058.
 PR
 XX 07-APR-2000; 2000DE-01019173.
 PR
 XX 30-JUN-2000; 2000DE-01032529.
 PR
 XX 01-SEP-2000; 2000DE-01043826.
 PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-010922/01.
 DR
 XX New nucleic acid derived from chemically treated metastasis genes, useful
 PT for diagnosis of cancers by analysis of cytosine methylation, also for
 PT treatment.
 PT
 XX Claim 1; SEQ ID NO 177; 23pp + Sequence Listing; English.
 PS The present invention provides a number of human metastasis associated
 XX genes which are modified by cytosine methylation. The sequences can be
 CC used in the diagnosis and treatment of cancer. The present sequence is
 CC one of the genes of the invention
 CC
 XX Sequence 15518 BP; 3047 A; 362 C; 4183 G; 7926 T; 0 U; 0 Other;
 SQ
 Query Match 7.0%; Score 62.8; DB 6; Length 15518;
 Best Local Similarity 46.6%; Pred. NO. 0.0032;
 Matches 234; Conservative 0; Mismatches 267; Indels 1; Gaps 1;
 QY 34 AATTAAATTCAAAAAATATTTGTAATAAAAAACCGACCAAGTTGATCGGTATTTAAT 93
 Db 942 AACTACGACCTAAATAAATAAAAAACAACTTCAAAATCTATTATAAAATTTCTTT 883
 QY 94 TATGTATAAAGATTCACATCTCTGGATCGAACCGGGGTCTGTACTATGCGCAAGATA 153
 Db 882 ACTAAAAAATAAAAAAATTTCTAAACAAATCAACACATATTAAATTTAATAAATA 823
 QY 154 CTATCTACCTAGACCATTTGTTTAAAGACTGTCTTTTAAAGACTGTCTTTTAAATTTAT 213
 Db 822 AAAAATACTATTAAAAATTAATACTAATACTCTCCATCTATTCTATATTTCTTAAAT 763
 QY 214 ACTCTTTAATTAATTTTGCACGAAATAAACCGACCAAGTTGGTTCGATTTTATAAAA 273
 Db 762 AATTATATCTTAA 704
 QY 274 AGTAAATTTACTTACCAAAATTAATACTTAAATTAATACTTAAATTAATACTAATA 333
 Db 703 AAAATCTTTCAAAAATAAATACTTAAATTAATACTTAAATTAATACTAATA 644
 QY 334 AATTTCGTAGTTTCTTTTAAATTAATACTTAAATTAATACTTAAATTAATACTAATA 584
 Db 643 CGGTAAATTTCTTTTAAATTAATACTTAAATTAATACTTAAATTAATACTAATA 524
 QY 394 GTTAGTCGGTTTCTTTGAAACATATAATTTTCGGGGACTCAAAATAGTTTCCGGATTTT 453
 Db 583 CTCGCTAGCTTACCAAAATTAATACTTAAATTAATACTTAAATTAATACTAATA 524
 QY 454 GCGCCAAAGAAACCGACCAAGTTGTCGGTTTCTGTAATAAAAAAATAAAAAA 513
 Db 523 AACAAATAAATCTATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 464
 QY 514 TATATTTTAAACCGACCAA 535
 Db 463 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 464

Oy 514 TATATTTTAAAAACCGACCA 535
| | | | | | | | | |
Db 463 AAAAAAAAAAAAAAAAAAAAA 442

Search completed: November 11, 2004, 08:10:27
Job time : 448.342 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 07:44:14 ; Search time 2796.79 Seconds
(without alignments)
11739.223 Million cell updates/sec

Title: US-10-009-966C-1_COPY_2255_3155
Perfect score: 901
Sequence: 1 tggctggaattaccagacca.....atgacttcattcgtattcttc 901

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	75.4	8.4	1101	CNS0021J	AL061936 Drosophila
2	70.4	7.8	863	BZ088961	BZ088961 lla55e08.
3	67.2	7.5	1092	CNS020K7	AL175696 Tetradon
4	66.8	7.4	909	AZ670108	AZ670108 ENTMP51TF
5	66	7.3	605	CL539617	CL539617 OB_Ba005
6	65.8	7.3	759	CL799614	CL799614 OR_Cha001
7	65.8	7.3	760	CL829972	CL829972 OR_Cha005
8	65.8	7.3	1101	CNS00EVL	AL069706 Drosophila
9	65.6	7.3	642	CG962071	CG962071 MBDE79TR
10	65.6	7.3	905	CNS00KHX	AL077798 Drosophila
11	65.4	7.3	748	BZ413877	BZ413877 i518e02.g
12	65.4	7.3	818	CC714001	CC714001 CGWDL09TV
13	65.2	7.2	736	BZ755113	BZ755113 PUFBZ03TD
14	65.2	7.2	757	BZ991051	BZ991051 PUFBNH65TD
15	64.8	7.2	715	CC335052	CC335052 OGOAX64TV
16	64.8	7.2	1037	BQ648574	BQ648574 AGENCOURT
17	64.6	7.2	802	CG730035	CG730035 OGWDR16TM
18	64.6	7.2	919	CG442944	CG442944 OGFBJ87TV
19	64.4	7.1	836	CNS01100	AL093642 Drosophila
20	64.4	7.1	845	CC993538	CC993538 ZUAH175TV
21	64.4	7.1	856	CB624068	CB624068 OSI1Ea12E
22	64.2	7.1	1029	CNS01ZGM	AL174271 Tetradon
23	64	7.1	651	CG054601	CG054601 PUFQX71TD
24	64	7.1	653	CG054598	CG054598 PUFQX71TB

C	25	64	7.1	694	7	CR292249	CR292249
	26	64	7.1	770	8	CC375169	CC375169
	27	64	7.1	837	8	BZ976845	BZ976845
	28	64	7.1	840	8	CC443649	CC443649
C	29	64	7.1	898	8	CC443646	CC443646
	30	63.8	7.1	774	9	CG440592	CG440592
C	31	63.8	7.1	792	9	CG632269	CG632269
C	32	63.8	7.1	797	8	CC007378	CC007378
	33	63.8	7.1	891	9	CC623348	CC623348
C	34	63.8	7.1	913	9	CG214581	CG214581
	35	63.8	7.1	924	9	CG174581	CG174581
	36	63.6	7.1	469	4	B1678504	B1678504
	37	63.4	7.0	645	8	BZ814249	BZ814249
	38	63.4	7.0	709	8	BH941486	BH941486
C	39	63.4	7.0	930	9	CG308156	CG308156
	40	63	7.0	864	8	BZ744789	BZ744789
	41	62.8	7.0	915	8	CC370715	CC370715
	42	62.6	6.9	1101	9	CNS00LT2	AL078714 Drosophila
C	43	62.4	6.9	759	9	CNS060XV	AL411257 T7 end of
	44	62.4	6.9	1101	9	CNS004ZW	AL055440 Drosophila
C	45	62.2	5.9	876	8	BZ635267	BZ635267

ALIGNMENTS

RESULT 1
CNS0021J
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION
AL061936
VERSION
AL061936.1
KEYWORDS
GSS:
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR05N11"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 8.4%; Score 75.4; DB 9; Length 1101;
Best Local Similarity 39.9%; Pred. No. 5e-05;

Matches 273; Conservative 57; Mismatches 348; Indels 6; Gaps 2;

QY 34 AATTAAATTCAAAAAATATGTGAAAAAAGCCGACCAAGTGTATCGGTATTTAAT 93
 DB 375 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 434
 QY 94 TATGTATATAAAGATTCACTATCTGGATCGGAATCGAACCGGGCTGTACTATGCGAAGATA 153
 DB 435 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 494
 QY 154 CTATTCACCACTAGACATGGTTCATTTTGTGTTTAAAGCTGCTTTTATTTGATTAT 213
 DB 495 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 554
 QY 214 ACTCTTAAATATATTTTTCACGAAATAACCGACCAAGTGTGGTATTTTAAATA 273
 DB 555 TTTTAAATTTTAAATTTTAAAWAAGAAATTAATAAAGAAWTTTAAWTTTAAW 614
 QY 274 AGTAAATTAATCTTACCAAGTGTGCTGATTTTAAATGATCGCGCAATTAACCGACC 333
 DB 615 WAAAAAAGAAATTTTAAAWAATTTTAAATTAATTAATAAAGAAWATAAATTTT 674
 QY 334 AATTTCGTAGTTTTTAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAA 393
 DB 675 TTAATWATWAAATTTTAAAGAAWTTTAAATTTTAAATTTTAAATTTTAAAT 734
 QY 394 GTTAGTCGTTCTTGAACATAAATTTTCGGGACTCAAAATAGTTT-----CCGCGAT 449
 DB 735 TTTAAAWATWAWTTTAAATTTTAAATTAATTAATTAATTAATTAATTTTAAAT 794
 QY 450 TTTTCGCGCAAGAAAGCCGACCAAGTGTGCTGTTTCGTAAAGAAAAAATTTAA 509
 DB 795 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAA 854
 QY 510 AAAATATATTTTAAAGAACCAACTTAGTCTGCGTTTTTGGTGTGATTTTGGACCGA 569
 DB 855 TTTTAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 914
 QY 570 CCAAGTGTGCTGCGTGTGCTGCTGTTTTCGCGAAATTTCTAG--TAGTCACCGAAC 627
 DB 915 AAAAATGKAGATGKGGTARKKGGTATGATGATGATGATGATGATGATGATGAT 974
 QY 628 CCTGTAGCTTCGGGAGAAATTTTGTATATGATGATGATGATGATGATGATGAT 687
 DB 975 KKTAAAAAADKAGAAAAATAGAAAGKATGAAAAAGTGGKAAATATTTKGAAGAAKAAA 1034
 QY 588 TTAAGAACGNGCACCTGAATA 711
 DB 1035 AAAAAAAGAAAAAAGAAAGAAK 1058

RESULT 2
 LOCUS BZ088961 863 bp DNA linear GSS 10-OCT-2002
 DEFINITION lla55e08.b1 B.oleracea002 Brassica oleracea genomic, genomic survey sequence.
 ACCESSION BZ088961
 VERSION BZ088961.1 GI:23723478
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 863)
 AUTHORS Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T., Nash, W., Rabinowicz, P.D. and Wilson, R.K.
 TITLE Whole genome shotgun reads from Brassica oleracea
 JOURNAL Unpublished (2002)
 COMMENT Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu

Plate: lla55 row: e column: 08
 Seq primer: -21UPPOT forward
 Class: shotgun
 High quality sequence stop: 521.
 Location/Qualifiers
 1..863
 source

FEATURES
 source

/organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea002"
 /note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 7.8%; Score 70.4; DB 8; Length 863;
 Best Local Similarity 66.0%; Pred. No. 0.00051;
 Matches 101; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 73 AAAGTTGATCGGTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 132
 DB 357 AAAACAGTTTGACITGCAAAATCACAAGAAAAATGACACCGCGTAAATCGAACCG 416
 QY 133 GGTCTGTAATGGAAGATATCTTACCACTAGACCATGTTTCAATTTGTTTAAAG 192
 DB 417 GGTCTGTAATGGAAGATATCTTACCACTAGACCATGTTTCAATTTGTTTAAAG 476
 QY 193 ACTGCTTTTATTTGATTTATCTCTTTAAATTA 225
 DB 477 ATTCCTCTTCANATATACAAAGCTTAACATA 509

RESULT 3
 CNS020K7/c

LOCUS CNS020K7 1092 bp DNA linear GSS 01-SEP-2000
 DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 2281 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION AL175696
 VERSION AL175696.1 GI:7813753
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE
 AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizanes, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W., and Weissenbach, J.
 Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)

TITLE

JOURNAL MEDLINE
 PUBMED 20296633
 REFERENCE 10835645

AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fizanes, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)

TITLE

JOURNAL MEDLINE
 PUBMED 10899143
 REFERENCE 3 (bases 1 to 1092)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

Plate: 0051 row: K column: 04
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES

Location/Qualifiers
1..605
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_Ba0051K04"
/tissue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_lib="OB_Ba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 7.3%; Score 66; DB 9; Length 605;
Best Local Similarity 73.7%; Pred. No. 0.0039;
Matches 84; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 100 ATAAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGCGAAGATACTATTC 159
Db 488 AAAAAGAAATCCACGACCGGGAATCGAACCGGGTCTGTACTATGCGGAGTACTATTC 547
QY 160 TACCCTAGACCATGCTTCATTTGTTTAAAGACTGCTTTTATTTGATTTAT 213
Db 548 TACCCTAGACCATGCTTCATTTGTTTAAAGACTGCTTTTATTTGATTTAT 601

RESULT 6
CL799614
LOCUS
DEFINITION OR_CBA0010001.f OR_CBA Oryza rufipogon genomic clone OR_CBA0010001
5', genomic survey sequence.
ACCESSION CL799614
VERSION CL799614.1 GI:51023683
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon

REFERENCE
AUTHORS Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0010 row: O column: 01
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
Location/Qualifiers
1..759
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBA0010001"
/tissue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_Cba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
drk treated 36 hrs before harvest"

FEATURES

Location/Qualifiers
1..759
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBA0010001"
/tissue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_Cba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
drk treated 36 hrs before harvest"

ORIGIN

Query Match 7.3%; Score 65.8; DB 9; Length 759;
Best Local Similarity 66.7%; Pred. No. 0.0042;
Matches 94; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 104 AAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGCGAAGATACTATTC 163
Db 526 AAAAATGCCACGACCGGGAATCGAACCGGGTCTGTACTATGCGGAGTACTATTC 585
QY 164 ACTAGACCATGCTTCATTTGTTTAAAGACTGCTTTTATTTGATTTATTTAAT 223
Db 586 ACTAGACCATGCTTCATTTGTTTAAAGACTGCTTTTATTTGATTTATTTAAT 645
QY 224 TATATTTTGCACGAAATAA 244
Db 646 TTCCCTGCAAGCGCAACAA 666

RESULT 7

CL829972
LOCUS
DEFINITION OR_CBA0051N11.f OR_CBA Oryza rufipogon genomic clone OR_CBA0051N11
5', genomic survey sequence.
ACCESSION CL829972
VERSION CL829972.1 GI:51075582
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon

REFERENCE
AUTHORS Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0051 row: N column: 11
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
Location/Qualifiers
1..760
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBA0051N11"
/tissue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_Cba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
drk treated 36 hrs before harvest"

FEATURES

Location/Qualifiers
1..760
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBA0051N11"
/tissue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_Cba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
drk treated 36 hrs before harvest"

ORIGIN

Query Match 7.3%; Score 65.8; DB 9; Length 760;
Best Local Similarity 66.7%; Pred. No. 0.0042;
Matches 94; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 104 AAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGCGAAGATACTATTC 163
Db 526 AAAAATGCCACGACCGGGAATCGAACCGGGTCTGTACTATGCGGAGTACTATTC 585

```

QY 164 ACTAGACCATTGGTCAATTTGTTTAAAGACTGCTTTTATTGATTGATTACTCTTTAAT 223
Db 586 ACTAGACCATTGGTGGTTCGTTGTTAGTGAGACAATTTCTCAATTAATAGTCTATAT 645

QY 224 TATATTTTTCACGAAATAA 244
Db 646 TTCCTGCAAGCGCAACAA 666

RESULT 8
CNS00EVL/c 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly) genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
Genoscope.
Direct Submission
AUTHORS Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
Location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/clone="end : T7"

ORIGIN
Query Match 7.3%; Score 65.8; DB 9; Length 1101;
Best Local Similarity 32.4%; Pred. No. 0.0041;
Matches 188; Conservative 105; Mismatches 284; Indels 4; Gaps 1;

QY 21 ACTTTGGTCGGTCAATTAATCAAAAAAATATTTGTAATAAAAAACCGCAAGTTGA 80
Db 1018 AAWTTAATAATATATATWATWATATATATATAWATAWATAWATAWATAWATAWAT 959

QY 81 TCGGTATTTTATGTATATAAAGATTCACACTCTCTGGAATCGAACCGGGTCTGTA 140
Db 958 AAWTTTAAATAAATAATWAWWTTTATTTTATTTTATTTTATTTTATTTTATTTAT 899

QY 141 CTATGGCAAGTACTATCTACCACTAGACCATTTGGT---TCATTTTGGTTTAAAGACTG 196
Db 898 WTWAAAWAAWAAWAAWTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 839

QY 197 TCTTTTATTCATTTACTCTTTAAATATATTTTTCACAGAAATAACCGACCAAGTT 256
Db 838 TATWATWATATWTTTWTATWATATTTATTTATTTATTTATTTATTTATTTATTTAT 779

```

```

QY 257 GGTGCAATTTTATTAATAAGTAAATTAATCTACCAAGTTGGTCGATTTTAAATGATC 316
Db 778 WATWWTWATAWATAATATATWATATTTATTTATTTATTTATTTATTTATTTATTTAT 719

QY 317 CGCGCAATTAACCGACCAATTTTGGTAGTGTTTTAAATATAATTTTATTTATTTAA 376
Db 718 ATWATTTTWTATTTTATTTTATTTTATTTTAAATTTAAATTAWAAATWATTT 659

QY 377 TTGAAAAACTAACCAAGTTAGTCGGTTTCTTGAAACAATAATTTTCGGCGGACTCAAAA 436
Db 658 ATTAAATAAATAATTAATAATTTTAAATAATTTTAAATAATTTTATTTATTTATTA 599

QY 437 TAGTTTCCCGCATTTTTCGCCCAAGAAACCGACCAAGTTGGTCGGTTTCGTAATAAA 496
Db 598 WATTTTATAWATTAATATATATWATTTTAAATTTTATTTTATTTTAAATTTATTTT 539

QY 497 AAAAAAATTTTAAATAATATTTTAAAAAAACCGACCAACTTTAGTCGGTTTTCGGTCG 556
Db 538 TTATWATAWATTTAAATAATWAAWAAATWAAWAAATTTTATTTTATTTATTTATTA 479

QY 557 ATTTTTCACCGACCAAGTTGGTCGGTCGACCTTTGGTCGG 597
Db 478 WWAATTTWKKKKKANADTGAGARATKDKKKKGGSGS 438

RESULT 9
CG962071 642 bp DNA linear GSS 15-DEC-2003
LOCUS MBEDE7STRC mch2 Medicago truncatula genomic clone 31M13, genomic
DEFINITION survey sequence.
ACCESSION CG962071
VERSION CG962071.1 GI:39884351
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 642)
TOWN, C.D., Shetty, J., Koo, H. and Feldblyum, T.F.
AUTHORS Sequencing of BAC ends from Medicago truncatula
TITLE Unpublished (2003)
JOURNAL Other GSSs: MBEDE79TFC
COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
Seq primer: CAGGAACAGCTATGACC
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1. .642
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="genotype Al7"
/db_xref="taxon:3880"
/clone="31M13"
/clone_lib="mth2"
/notes="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Cook, D.R. and Kim, D.J, unpublished"

ORIGIN
Query Match 7.3%; Score 65.6; DB 9; Length 642;
Best Local Similarity 76.9%; Pred. No. 0.0047;
Matches 80; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 99 AATAAAAGATTCACATCTCTGGAATCGAACCGGGTCTGTACTATGGCAAGACTATT 158
Db 535 AAAAAAATTTGCACCGCGGGAATCGAACCGGGTCTGTACTATTTGGCAGGCTACTATT 594

```

```

QY 159 CTACCACTAGACATTGGTTCATTTGTTTAAAGACTGTCCTTT 202
Db 159 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
595 CTACCACTAGACACTGGTGTGTTGATGTTAAGAGGTTTAT 638
595 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
CNS000KHx/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC;
BACR17N06 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL077798
VERSION
AL077798.1 GI:4957174
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 905)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazuhiro Osogawa and
Aaron Mamoser in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCL-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
Location/Qualifiers
1..905
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR17N06"
/clone_lib="RPCL-98"
/note="end : T7"

ORIGIN
Query Match 7.3%; Score 65.6; DB 9; Length 905;
Best Local Similarity 34.5%; Pred. No. 0.0046;
Matches 134; Conservative 71; Mismatches 183; Indels 0; Gaps 0;

QY 148 AAGATACCTATCTACCACTAGACATTGGTTCATTTGTTTAAAGACTGTCCTTTTATTTG 207
Db 148 AAAAAAAAAAAWAAAAAABAAAATTTATATATTTTATTTTATTTTATTTTATTTT 839

QY 208 ATTATATCTCTTAAATATATTTTTCACGAAATATACCGACCAAGTTGTCGATTTTA 267
Db 208 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
779 TATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 779

QY 268 TTAAGAGTAAATATCTACCAAGTTGGTCATTTTATTTTAAATGATCCGCGAATTA 327
Db 268 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 WATATATATCTCTTAAATATATTTTTCACGAAATATACCGACCAAGTTGTCGATTTTA 719

QY 328 CCGACCAATTTTGTAGTTTATTTTAAATATATTTTATTTTATTTTATTTTAAATGAAACTA 387
Db 328 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 659

QY 388 ACCAAGTGTAGTCTTCTTGAACATATATTTTCGGGGACTCAAAATAGTTTCCCGC 447
Db 388 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
658 TTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 599

```

```

QY 448 ATTTTGGCCCAAGAAACCGACCAAGTTGGTTCGTTTAAAGAAAAAATAATTT 507
Db 448 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
598 DTTTTTTTAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 539
598 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 508 AAAAAATATATTTTAAAAAACCGACCAA 535
Db 508 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
538 MWMCMCAACATTATTAAMAMCAAWTAA 511
538 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
BZ413877/c
LOCUS
DEFINITION
BZ413877
ACCESSION
BZ413877.1 GI:26355893
VERSION
BZ413877
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 748)
Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
Katzemburger,F., King,L., Miller,S., Muller,S., Nascimento,L.,
Zutavern,T., McCombie,W.R. and Martienssen,R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Auerberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: if18 row: e column: 02
Seq primer: -21M31univrev
Class: shotgun
High quality sequence stop: 748.
Location/Qualifiers
1..748
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/clone="if18e02"
/clone_lib="DH5a"
/lab_host="DH5a"
/clone_lib="WGS-ZmaysF (DH5a methyl filtered)"
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector (.x/y
reads in M13mp19, .b/g reads in pUC19). The same ligation
was transformed into DH5a."

ORIGIN
Query Match 7.3%; Score 65.4; DB 8; Length 748;
Best Local Similarity 63.2%; Pred. No. 0.0051;
Matches 117; Conservative 0; Mismatches 66; Indels 2; Gaps 1;

QY 53 ATTGTAAAAAACCACCGACCAAGTTGATCGTATTTTAAATGATTAATAAAGA--TT 110
Db 450 ATTACACATATGATAGAAAAAATAATGAAAAAAGAAAAAATAAATAAATAATG 391

QY 111 CACTATCTGGGAATCGAACCGGGGTCTGTACTATGGCAAGATACATTTCTACCACTAGAC 170
Db 390 CACCAGCGGGAATCGAACCGGGGTCTGTACTATGGCAAGATACATTTCTACCACTAGAC 331

QY 171 CATTTGGTTTCATTTTGTATAGACTGCTTTTATTTGATTTATATCTCTTTAATATATT 230
Db 171 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

Db      330  CACTGGTGCAATTTGTGCTTAAGTTCTTAATAGGTTGATTATAGGACCGTGCTACATGT 271
QY      231  TTGCA 235
Db      270  TCACA 266

RESULT 12
CC714001/c
LOCUS
DEFINITION
  CC714001 818 bp DNA linear GSS 19-JUN-2003
  OGDWLO9TV ZM 0.7-1.5.KB Zea mays genomic clone ZMMBma0547A18,
  genomic survey sequence.
ACCESSION
  CC714001
VERSION
  CC714001.1 GI:32118777
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 818)
REFERENCE
  AUTHORS
    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
    Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
    Citek,R.W., Nunberg,A., Robbins,D. and Laxey,N.
  TITLE
    Consortium for Maize Genomics
  JOURNAL
    Unpublished (2002)
  COMMENT
    Other GSSs: OGDWLO9TH
    Contact: Cathy Whitelaw
  TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 301-838-5843
    Fax: 301-838-0208
    Email: whitelaw@tigr.org
    Seq primer: TF
    Class: sheared ends.
FEATURES
  source
    Location/Qualifiers
      1..818
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMMBma0547A18"
        /clone_lib="ZM 0.7-1.5 KB"
        /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
        methylation filtered genomic DNA library"
ORIGIN
  Query Match 7.3%; Score 65.4; DB 9; Length 818;
  Best Local Similarity 63.2%; Pred. No. 0.0051;
  Matches 117; Conservative 0; Mismatches 66; Indels 2; Gaps 1;

QY      53  ATTGTAAAAAACCACCAAGTTGATCGGTATTTTAATTATGTAATAAAAGA--TT 110
Db      489  ATTAAACAACATGATGAAAAAATAATGGGAAAAAAGAATAATGAATAAAAAATTG 430
QY      111  CACTATCTGGGAATCGAACCGGGTCTGTACTATGCGAAGATACTATTCTACCACTAGAC 170
Db      429  CACCAGCCGGGAATCGAACCGGGTCTGTACTACCGTGCAGGGTACTATTCTACCACTAGAC 370
QY      171  CATTGGTCTATTTTGTTTTAAGACTCTCTTTATTATTGATTATCTCTTTAATTATTT 230
Db      369  CACTGTGCAATTTGTGCTTAAGTTCTTAATAGGTTTGATTATAGGACCGTGCTACATGT 310
QY      231  TTGCA 235
Db      309  TCACA 305

RESULT 13
BZ755113/c
LOCUS
DEFINITION
  BZ755113 736 bp DNA linear GSS 10-MAR-2003
  PUFB203TD ZM 0.6-1.0.KB Zea mays genomic clone ZMMBta284B06,
  genomic survey sequence.

```

```

ACCESSION
  BZ755113
VERSION
  BZ755113.1 GI:28907462
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 736)
REFERENCE
  AUTHORS
    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
    Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
    Bennetzen,J.
  TITLE
    Maize Genomics Consortium
  JOURNAL
    Unpublished (2003)
  COMMENT
    Contact: Cathy Whitelaw
  TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 301-838-5843
    Fax: 301-838-0208
    Email: whitelaw@tigr.org
    Seq primer: TF
    Class: sheared ends.
FEATURES
  source
    Location/Qualifiers
      1..736
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMMBta284B06"
        /clone_lib="ZM 0.6-1.0 KB"
        /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
        Cot selected genomic DNA library"
ORIGIN
  Query Match 7.2%; Score 65.2; DB 8; Length 736;
  Best Local Similarity 67.9%; Pred. No. 0.0056;
  Matches 91; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY      99  AATAAAGATTCTACTATCTGGGAATCGAACCGGGTCTGTACTATGCGAAGATACTATT 158
Db      403  AATTAAAAATTGCACCGCCGGGATCGAACCGGGTCTGTACCGTGCAGGGTACTATT 344
QY      159  CTACACATAGACCATGTTGTTTCACTTTTGTTTTAAGACTGTCTTTTATTGATTATCTCT 218
Db      343  CTACCACTAGACCATGTTGTTTGTGTTTGTCTTTTAAATAGTTTAAATTATCTGT 284
QY      219  TTAATTATATTTT 232
Db      283  TCATCAATACATGT 270

RESULT 14
BZ991051/c
LOCUS
DEFINITION
  BZ991051 757 bp DNA linear GSS 25-MAR-2003
  PUFH65TD ZM 0.5-1.0.KB Zea mays genomic clone ZMMBta094K10,
  genomic survey sequence.
ACCESSION
  BZ991051
VERSION
  BZ991051.1 GI:29232461
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 757)
REFERENCE
  AUTHORS
    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
    Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
    Bennetzen,J.
  TITLE
    Maize Genomics Consortium
  JOURNAL
    Unpublished (2003)
  COMMENT
    Contact: Cathy Whitelaw
  TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA

```

Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

FEATURES

Location/Qualifiers
 1..757
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBta094K10"
 /clone_lib="ZM 0.6-1.0 KB"
 /note="vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
 Cor selected genomic DNA library"

ORIGIN

Query Match 7.2%; Score 65.2; DB 8; Length 757;
 Best Local Similarity 67.9%; Pred. No. 0.0056;
 Matches 91; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
 Qy 99 AATAAAAGATTCTACTATCTGGGAATCGAACCGGGGTCTGTACTATGGCAAGATCTATT 158
 Db 267 AATTAATAAATTGCACACCGCGGAATCGAACCGGGGTCTGTACCGTGGCAGGCTACTATT 208
 Qy 159 CTACCACTAGACCACTGGTTCATTTGTTTAAAGACTGCTCTTTTATTGATTATCTCT 218
 Db 207 CTACCACTAGACCACTGGTTCATTTGTTTAAAGACTGCTCTTTTAAATAGTTTAAATATCTGT 148
 Qy 219 TTAATTATATTTT 232
 Db 147 TCATCAATACATGT 134

RESULT 15

CC335052 715 bp DNA linear GSS 16-MAY-2003
 LOCUS OGOAX64TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0341L08,
 DEFINITION genomic survey sequence.

ACCSSION CC335052

VERSION CC335052.1 GI:30804465

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 715)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers
 1..715

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0341L08"

/clone_lib="ZM 0.7-1.5 KB"

/note="vector: pCISK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match 7.2%; Score 64.8; DB 8; Length 715;
 Best Local Similarity 66.4%; Pred. No. 0.0068;
 Matches 93; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 Qy 99 AATAAAAGATTCTACTATCTGGGAATCGAACCGGGGTCTGTACTATGGCAAGATCTATT 158
 Db 230 AATAAAAACTGCACACCGCGGAATCGAACCGGGGTCTGTACCGTGGCAGGCTACTATT 289
 Qy 159 CTACCACTAGACCACTGGTTCATTTGTTTAAAGACTGCTCTTTTATTGATTATCTCT 218
 Db 290 CTACCACTAGACCACTGGTTCATTTGTTTAAAGACTGCTCTTTTATTGATTATCTCT 349
 Qy 219 TTAATTATATTTTGCACGA 238
 Db 350 AATGTGACAGATTGGATGGA 369

Search completed: November 11, 2004, 12:05:39
 Job time : 2800.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 07:45:44 ; Search time 92.8475 Seconds
(without alignments)
6897.553 Million cell updates/sec

Title: US-10-009-966C-1_COPY_2255_3155
Perfect score: 901
Sequence: 1 tgcgcgaattaccagacca.....atgacttcattcgtattctc 901

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.6	5.5	3454	4	US-09-270-767-11909
2	48.6	5.4	640681	4	US-09-790-988-1
3	48.4	5.4	216	4	US-09-248-796A-7073
4	46.8	5.2	2435	3	US-09-306-593-1
5	45.8	5.1	660	1	US-07-991-867B-32
6	45.8	5.1	660	1	US-08-107-755A-32
7	45.8	5.1	660	2	US-08-544-332-32
8	45.8	5.1	660	4	US-09-370-861A-32
9	45.8	5.1	1511	1	US-07-991-867B-8
10	45.8	5.1	1511	1	US-08-107-755A-8
11	45.8	5.1	1511	2	US-08-544-332-8
12	45.8	5.1	1511	4	US-09-370-861A-8
13	45.8	5.1	4810	3	US-08-852-629-11
14	45.8	5.1	4838	3	US-08-852-629-15
15	45.8	5.1	11131	4	US-10-204-708-28
16	45.6	5.1	718	3	US-08-998-416-673
17	45	5.0	1099	4	US-09-270-767-11371
18	45	5.0	1141	4	US-09-806-708B-22
19	45	5.0	6040	4	US-10-204-708-69
20	43.8	4.9	1141	4	US-09-806-708B-22
21	43.8	4.9	29604	3	US-08-781-891-207
22	43.8	4.9	29604	4	US-09-618-166-207
23	43.6	4.8	13830	4	US-09-614-981-8
24	43	4.8	1055	4	US-09-806-708B-23
25	43	4.8	3348	4	US-09-976-594-363
26	42.6	4.7	1024	4	US-09-328-475C-81
27	42.6	4.7	3212	3	US-08-673-814-1

Sequence 1, Appli
Sequence 7, Appli
Sequence 13, Appli
Sequence 1, Appli
Sequence 93, Appli
Sequence 73, Appli
Sequence 71, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 43, Appli
Sequence 1, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-270-767-11909
; Sequence 11909, Application US/03270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11909
; LENGTH: 3454
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11909

Query Match 5.5%; Score 49.6; DB 4; Length 3454;
Best Local Similarity 46.5%; Pred. No. 0.028;
Matches 160; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 34 ANTTAAATCAAAAAATATGTAAAAAAGCCGACCAAGTGTGCGTATTTTAAAT 93
DB 492 AAAAAAAGAAAAAAGCAATCAATAAAGCAATTCGTCAATTTTAAATTC 551
QY 94 TATGTAATAAAGATTCACACTCTGGAATCGAACCGGGGTCTGTACTATGCGCAAGATA 153
DB 552 AGCATATATACATATATAATTTTCTTATGTTTACATACACACCAATTCGTAATCG 611
QY 154 CTAATCTACACTAGACCAATTCGTTCAATTTGTTTAAAGACTGCTTTTATTTGATTTAT 213
DB 612 CAAATGTATATAAATACCTCTTATATATCGTATATCTTTGAATTTGAAATTTAGCAGA 671
QY 214 ACTCTTTAATATATTTTTCGACGAAATAAAGCCGAAAGTGTGTCGATTTTATTTAA 273
DB 672 ACCTTGAATGAACCTTTTGTCTATATATATCTATCAATGTTGTCATGCGTGCAC 731
QY 274 AGTAAATTAATCAACCAAGTTGTCGATTTTAAATGATCCGCGAATTAACCGACC 333
DB 732 AATTTGACTGTATACCAATTTATATATATATATATATATATATATATATATATAT 791
QY 334 AATTTTGTAGTTTATTTTATATATATATATATATATATATATATATATATAT 377
DB 792 ACTTTGAGTAAATTTGTTTACAAAGAAATTCATGTTTCTCAAT 835

RESULT 2
US-09-790-988-1/c
; Sequence 1, Application US/09790988

```

Db      209  AACACAAAATGCAAGAACCGGGAAATCGAACCGGGGGCCCAACGATGCGCAACCTTGGATT 150
Qy      159  CTACCACTAGACCAATGGTTCATTTTGTTTTAAGACTGCTCTTTATTTGATTTTATACTCT 218
Db      149  TTACCACTAAACCAATTTCTTGCACCGCGCTAATGAAGTCTTTGTGTAAATTTTGGCCC 90
Qy      219  TT 220
Db      89  TT 88

RESULT 4
US-09-306-593-1/c
; Sequence 1, Application US/09306593
; Patent No. 6184018
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; APPLICANT: Ximenes, Eduardo A.
; TITLE OF INVENTION: Beta-glucosidase Coding Sequences and Protein from
; TITLE OF INVENTION: Orpinomyces PC-2
; FILE REFERENCES: 31-98us
; CURRENT APPLICATION NUMBER: US/09/306,593
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/084,494
; EARLIER FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2435
; TYPE: DNA
; ORGANISM: Orpinomyces sp. PC-2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(2009)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (87)..(2009)
US-09-306-593-1

Query Match 5.2%; Score 46.8; DB 3; Length 2435;
Best Local Similarity 52.9%; Pred. No. 0.11;
Matches 146; Conservative 0; Mismatches 127; Indels 3; Gaps 2;

Qy      148  AAGATACTATTCCACCTAGACCAATGGTCTCATTTTGTTTTAAAGACTGCTTTTATTTG 207
Db      2334  AACACAAAGCTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 2275
Qy      208  ATTTATACITCTTTAAATATATTTTTCGCAGAAATTAACCGACCAAGTTGGTCGATTTTA 267
Db      2274  ATTTGTTTCTCGCTTATATTTTATTTCTTAAAAAATATAAAATCTCTTTAGAAATA-TTTA 2216
Qy      268  TTAAGAAGTAAAATTACTTTACCAAGTTGGTCGATTTTTTTAAATGATCCGCGAATTAA 327
Db      2215  ATTAATATCATTTATTTCATTAATTTGTATATTATTTATTTATTTTATTTATGATTC 2156
Qy      328  CCGACCAATTTGGTAGG--TTTTTTAAATATAATTTTATTTTATTTTATTTTAAATGAAAAAC 385
Db      2155  AAAAAGATTTATCTATATAAATTTATATAATTAATAAATTAATTTATTTTCTACTACTATAG 2096
Qy      386  TAACCAAGTACTCGGTTCTTGAACACATAAATTT 421
Db      2095  CAAGAAGTGAATAAATAATTCATTATAAT 2060

RESULT 5
US-07-991-867B-32
; Sequence 32, Application US/07991867B
; Patent No. 5476781
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.

```

```

; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,867B
; FILING DATE: 12-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF114.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-07-991-867B-32

Query Match          5.1%; Score 45.8; DB 1; Length 660;
Best Local Similarity 44.4%; Pred. No. 0.14;
Matches 226; Conservative 0; Mismatches 282; Indels 1; Gaps 1;

QY 9 AATTACCGACCAACTTTGGTCGGTCAATTAATAAATTCAGAAATTTGGTAACCTTAATATGTTT 430
DB 12 AATTCTGATATAATTTTAATAATTTTAATAATTTTAATAATTTTAATAATTTTAATAATTTA 71
QY 69 GACCAAGTTGATCGGTATTTTAATTAATTAATAAAGATTCACCTATCTGGGAATCGAA 128
DB 72 AAACAGAGATAATGTTAATTAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 131
QY 129 CCGGGGCTCTACTATGCGCAAGATCACTATCTACCACTAGACCATTTGGTTCATTTTGT 188
DB 132 GCATATAATATTAAGTAATA-AATTTTAATAATATTCAGAAATTTAAAGTT 190
QY 189 TAGACTGCTCTTTTATTTGATTTTATCTCTTTTAAATTAATTTTGGCAGAAATAACCGA 248
DB 191 TATATATTTTCAAAATTTAAATTTAATTTAAATTTTAAATTTTAAACAAATTTAAATAAT 250
QY 249 CCAAAGTTGGTCGATTTTATTAATAAAGTAAATTTACTTACCAAGTTGGTCGATTTT 308
DB 251 CATATTTAGATATATCTTATAACAAATAAGCAATATAGTATATATTAATTAATTAAT 310
QY 309 AAATGATCCCGCAATTAACCGCAATTTTGGTAGGTTTTTTTAAATTAATTAATTTTAT 368
DB 311 CTATAGAAATTTTAAATTTGGAATCATGTAATAATAATGACTATAATTTTATTAATAAT 370
QY 369 TATTTAATTTGAAAACTAACCAAGTTAGTCGGTTCTTTGAAACATAAATTTTCGGGGA 428

```

```

DB 371 TAGTAAATTTAAAAAATAATAATAATTAATAATTAATAATTTGGTAACCTTAATATGTTT 430
QY 429 CTCRAAAATAGTTTCCCGCATTTTTCGCCCAAGAAACCCAGCAACCAAGTTGTCGGTTTC 488
DB 431 TTCCTATTAGTATAGTTGAGTTAAATATGGAATCAATACAAATTAAGATTATAATTA 490
QY 489 GTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 517
DB 491 TAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 519

RESULT 6
US-08-107-755A-32
; Sequence 32, Application US/08107755A
; Patent No. 5721352
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/107,755A
; FILING DATE: 19-AUG-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,658
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF114.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-107-755A-32

Query Match          5.1%; Score 45.8; DB 1; Length 660;
Best Local Similarity 44.4%; Pred. No. 0.14;
Matches 226; Conservative 0; Mismatches 282; Indels 1; Gaps 1;

QY 9 AATTACCGACCAACTTTGGTCGGTCAATTAATAAATTCAGAAATTTGGTAACCTTAATATGTTT 68
DB 12 AATTCTGATATAATTTTAATAATTTTAATAATTTTAATAATTTTAATAATTTTAATAATTTA 71
QY 69 GACCAAGTTGATCGGTATTTTAATTAATTAATAAAGATTCACCTATCTGGGAATCGAA 128
DB 72 AAACAGAGATAATGTTAATTAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 131
QY 129 CCGGGGCTCTACTATGCGCAAGATCACTATCTACCACTAGACCATTTGGTTCATTTTGT 188

```

```

Db      132 GCATATATATATATGATGATAATA-ATATTTTAAATAATAATCCAGAAATATTTAAAGTT 190
Qy      189 TAAGACTGCTCTTTTATTTGATTTTATACTCTTTTAAATATATTTTGCACGAAATAACCGA 248
Db      191 TATATATTTCAAAATTTAAATATATTTAAATTTTATAACAAAATTTAAAAATATATA 250
Qy      249 CCAAGTGTGCGATTTTATTAATAAGTAAATCTTACCAAGTGTGCGATTTT 308
Db      251 CATATTTAGATATATCTTATAACAAAATAGCAATATAGTAATATATATCTACCCATT 310
Qy      309 AAATGATCCGCGAATTAACCGCAATTTTGGTAGGTTTTTAAATATTAATTTTATT 368
Db      311 CTATAGAAATTTTAAATTTGATCATGTAATATATATGACTATATTTTATTAATTT 370
Qy      369 TATTTTAAATGAAAACTAACCAAGTTAGTCGGTTTCTTGAACATATAATTTCCGCGGA 428
Db      371 TAGTAAATTTTAAAAAATTAATAATCTTAAAAATTAATTTGGTAACTTTTAAATATGTTT 430
Qy      429 CTCAAAATAGTTTCCCGCATTTTTCGCGCAAGAAACCGACCAAGTTTGGTCGGTTTC 488
Db      431 TTCCTATTAGTATAGTTAGTTAAATATGGAATCAATACAAATTAAGATTAATAATTTA 490
Qy      489 GTAAAAAATTAATTTTAAAAATATA 517
Db      491 TAGAAAAATTAATTTTAAAAAATTA 519

```

RESULT 7

```

US-08-544-332-32
; Sequence 32, Application US/08544332
; Patent No. 5935777
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gerard H. Bencen
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/544,332
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,867
; FILING DATE: 07-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/107,755
; FILING DATE: 19-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H.
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: UF114.C4

```

TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-544-332-32

```

Query Match 5.1%; Score 45.8; DB 2; Length 660;

Best Local Similarity 44.4%; Pred. No. 0.14;

Matches 226; Conservative 0; Mismatches 282; Indels 1; Gaps 1;

```

Qy      9 AATTAACCGACCAACTTTGGTCGGTCAATTAATTTCAAAAAAATATTTGTAATAAAAAAC 68
Db      12 AAATTTCTGATATAATTTTAAATAATATTTTAAAAATATATATATTTAAAAAATAAT 71
Qy      69 GACCAAGTGTGATCGTATTTTAAATATATTAATAAAAAAGATTCACATCTCGGAATCGAA 128
Db      72 AACACAGATATGTTATTAATTAATTAATTAATAAAAAATTTAGTTAATTAGAGATT 131
Qy      129 CCGGGTCTGTACTATGCGAAGATCTATTTCTACACTAGACCAATTTGGTTCATTTTGT 188
Db      132 GCATATAATATATATGATAATA-ATATTTTAAATAATATTTCCAGAAATATTTAAAAAGTT 190
Qy      189 TAAGACTGCTCTTTTATTTGATTTTACTCTTTTAAATATATTTTGCACGAAAAATACCGA 248
Db      191 TATATATTTCAAAATTTAAATATTTTAAATTTTAAATTTTAAACAAATTTAAAAATATAA 250
Qy      249 CCAAGTGTGCGATTTTATTTAAAAAGTAAATTTACTTACCAAGTTGGTCGATTTTTT 308
Db      251 CATATTTAGATATATCTTTATAACAAAATAGCAATATAAGTAATATATATATATACCAATT 310
Qy      309 AAATGATCCGCGAATTAACCGCAATTTTGGTAGGTTTTTAAATATTAATTTTATTTAT 368
Db      311 CTATAGAAATTTTAAATTTGTAATCATGTAATAATAAATGACTATAATTTTATTAATTT 370
Qy      369 TATTTTAAATGAAAACTAACCAAGTTAGTCGGTTTCTTGAACATATAATTTTCGCGGA 428
Db      371 TAGTAAATTTTAAAAAATTAATAATCTTAAAAATTAATTTGGTAACTTTTAAATATGTTT 430
Qy      429 CTCAAAATAGTTTCCCGCATTTTTCGCGCAAGAAACCGACCAAGTTTGGTCGGTTTC 488
Db      431 TTCCTATTAGTATAGTTAGTTAAATATGGAATCAATACAAATTAAGATTAATAATTTA 490
Qy      489 GTAAAAAATTAATTTTAAAAATATA 517
Db      491 TAGAAAAATTAATTTTAAAAAATTA 519

```

RESULT 8

```

US-09-370-861A-32
; Sequence 32, Application US/09370861A
; Patent No. 6410221
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 6410221e1 Entomopoxvirus Expression System
; FILE REFERENCE: UF114.C4.D1
; CURRENT APPLICATION NUMBER: US/09/370,861A
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 07/991,867
; PRIOR FILING DATE: 1992-12-07
; PRIOR APPLICATION NUMBER: US 08/107,755
; PRIOR FILING DATE: 1993-08-19
; PRIOR APPLICATION NUMBER: WO 92/14818
; PRIOR FILING DATE: 1992-02-12
; PRIOR APPLICATION NUMBER: US 07/827,685
; PRIOR FILING DATE: 1992-01-30

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
FILING DATE: 12-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UFL14.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-07-991-867B-8

Query Match 5.1%; Score 45.8; DB 1; Length 1511;
Best Local Similarity 44.4%; Pred No 0.17;
Matches 226; Conservative 0; Mismatches 282; Indels 1; Gaps 1;

Qy 9 AATTACCGACCAACTTGGTCGGTCAATTAAATTCAAAAAAAATATGTGTAACCAAAACC 68
Db 863 AAATTCGTATATAATTTTAATAAATATTTTAAATATTTATAATTTAAAAAATAATAAT 922
Qy 69 GACCAAGTGTATCGGTATTTTAAATGTAATAAAGAGATTCACATATCTGGGAATCGAA 128
Db 923 AAACAGAGATAATGTTATTAAATTAATATATTATAAATAAATAGTTAAATTTAGAAGAATT 982
Qy 129 CCGGGGTCTGTACTATGGCAAGTACTATTCTACCACATAGACCATTTGGTTCATTTTGT 188
Db 983 GCATATAATATATTATGATAATA-ATATTTTAAATAATTTCCAGAAAATATTAAAGATT 1041
Qy 189 TAGACATGCTTTTATTGATTTTATCTCTTTTAAATTAATTTTTTGCACGAAAATAACCGA 248
Db 1042 TATATATTTCAAATTTAAATTAATTAATTTAAATTTTATAACAAAATTTAAAAAATAAA 1101
Qy 249 CCAAGTGGTCGATTTTATTAAAAAGTAAAAATTACTTACCAAGTTGGTCGATTTTTTT 308
Db 1102 CATATTTAGATATATCTTATAACAAAAATAGCAATATAAGTATATTTATCTACCAATT 1161
Qy 309 AAATGATCCGCCGAATTAACCGCAATTTTGGTAGGTTTTTTTAAATTAATTTTATT 368
Db 1162 CTATAGAAATTTTAAATTTGGAATCATGTAAATAATAAATAGTACTATAATTTTATTAAAT 1221
Qy 369 TATTTTAAATGAAAACTAACCGCAATTTTGGTAGGTTTTTTTAAATTAATTTTATT 428
Db 1222 TAGTAAATTTAAAAAATAATTAATTAATCTTAAAAATAAATTTTGGTAACTTTTAATAATGTTT 1281
Qy 429 CTCAAAAATAGTTTCCCGCATTTTTTGGCGCCAAAAGAAAACCGACCAAGTTGGTCGGTTTC 488

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 10:27:15 ; Search time 482.239 Seconds
(without alignments)
10091.039 Million cell updates/sec

Title: US-10-009-966C-1_COPY_2255_3155
Perfect score: 901
Sequence: 1 tggctgaattaccgacca.....atgacttcattcgtattctc 901

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3625171 seqs, 2700493622 residues

Total number of hits satisfying chosen parameters: 7250342

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:
19: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:
20: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:
21: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	64.8	7.2	1140	US-10-190-312A-97	Sequence 97, Appl
C 2	64.8	7.2	1140	US-10-190-312A-119	Sequence 119, Appl
C 3	64.2	7.1	2131	US-10-424-599-109423	Sequence 109423,
C 4	64	7.1	5945	US-10-311-455-58	Sequence 58, Appl
C 5	62.8	7.0	15518	US-10-311-455-2145	Sequence 2145, Ap
C 6	62.8	7.0	15518	US-10-240-485-177	Sequence 177, Appl
C 7	62.4	6.9	2000	US-09-887-576-129	Sequence 129, Appl
C 8	58	6.4	3673778	US-10-312-841-1	Sequence 1, Appli
C 9	57.2	6.3	3673778	US-10-312-841-1	Sequence 1, Appli
C 10	56.2	6.2	552	US-10-021-323-93	Sequence 93, Appl
C 11	55.4	6.1	6239	US-10-311-455-1749	Sequence 1749, Ap
C 12	55.4	6.1	6239	US-10-240-453-171	Sequence 171, Appl

C 13	55.2	6.1	5407	15	US-10-311-455-2064	Sequence 2064, Ap
C 14	54.6	6.1	368	9	US-09-834-975-47	Sequence 47, Appl
C 15	53	5.9	6092	16	US-10-221-613-36	Sequence 36, Appl
C 16	52.8	5.9	8979	15	US-10-311-455-757	Sequence 757, App
C 17	52.8	5.9	8979	16	US-10-321-613-137	Sequence 137, App
C 18	52.6	5.8	520	17	US-10-021-323-7699	Sequence 7699, Ap
C 19	52.4	5.8	396	17	US-10-021-323-17801	Sequence 17801, A
C 20	52.4	5.8	419	9	US-09-980-352-11234	Sequence 11234, A
C 21	51	5.7	16914	16	US-10-221-613-214	Sequence 214, App
C 22	50.4	5.6	5756	15	US-10-311-455-559	Sequence 559, App
C 23	50.4	5.6	15732	14	US-10-239-676-96	Sequence 96, Appl
C 24	50.4	5.6	15732	15	US-10-240-453-108	Sequence 108, App
C 25	50.2	5.6	469	17	US-10-021-323-16830	Sequence 16830, A
C 26	50.2	5.6	500	10	US-09-814-353-5831	Sequence 5831, Ap
C 27	50.2	5.6	500	10	US-09-814-353-12112	Sequence 12112, A
C 28	50.2	5.6	520	17	US-10-021-323-7699	Sequence 7699, Ap
C 29	50	5.5	5997	15	US-10-311-455-1598	Sequence 1598, Ap
C 30	49.8	5.5	3673778	15	US-10-312-841-2	Sequence 2, Appli
C 31	49.6	5.5	5520	15	US-10-311-455-1491	Sequence 1491, Ap
C 32	49.4	5.5	361	18	US-10-425-115-138029	Sequence 138029,
C 33	49.2	5.5	499	17	US-10-021-323-10189	Sequence 10189, A
C 34	49	5.4	6963	15	US-10-311-455-952	Sequence 952, App
C 35	49	5.4	15046	15	US-10-311-455-2074	Sequence 2074, Ap
C 36	48.8	5.4	296	10	US-09-814-353-5940	Sequence 5940, Ap
C 37	48.8	5.4	296	10	US-09-814-353-12220	Sequence 12220, A
C 38	48.6	5.4	5867	17	US-10-433-793-181	Sequence 181, App
C 39	48.6	5.4	6052	16	US-10-221-613-352	Sequence 352, App
C 40	48.6	5.4	64681	9	US-09-790-988-1	Sequence 1, Appli
C 41	48.4	5.4	875	14	US-10-198-846-2872	Sequence 2872, Ap
C 42	48.4	5.4	17142	14	US-10-239-676-206	Sequence 206, App
C 43	48.4	5.4	17142	15	US-10-311-455-2080	Sequence 2080, Ap
C 44	48.4	5.4	17142	15	US-10-240-453-304	Sequence 304, App
C 45	48.2	5.3	5641	15	US-10-311-455-1369	Sequence 1369, Ap

ALIGNMENTS

RESULT 1

US-10-190-312A-97/c
; Sequence 97, Application US/10190312A
; Publication No. US2003019468A1
; GENERAL INFORMATION:
; APPLICANT: Chromagenics B.V.
; APPLICANT: Kruckeberg, Arthur L.
; TITLE OF INVENTION: DNA sequences comprising gene transcription regulatory qualities
; TITLE OF INVENTION: methods for detecting and using such DNA sequences
; FILE REFERENCE: 2183-4993.1
; CURRENT APPLICATION NUMBER: US/10/190,312A
; PRIOR FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: 60/303,199
; NUMBER OF SEQ ID NOS: 1079
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of STAR A13
US-10-190-312A-97

Query Match	7.2%	Score 64.8;	DB 15;	Length 1140;
Best Local Similarity	60.0%;	Pred. No. 0.0011;		
Mismatches 108;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0;
QY	98	TAATAAAAGATTCTACTTCTGGGATCGAACCGGGTCTGTACTATCGCAAGTACTAT	157	
Db	366	TCATAGCTCAATGACACCGCGGGAATCGAACCGGGTCTGTACCGTGGCGGTACTAT	307	
QY	158	TCATACACTAGACCAATGGTTTCATTTGTTTAAAGACTGCTTTTATTGATTATCTC	217	

Db 306 TCTACCACTAGACCACTGGTCTTTTGTGATTAACACGACTAAATATATATATATAGAT 247
 Qy 218 TTTAATTATATTTTGCACGAAATAACCGACCAAGCTTGCTGCAATTTTATTAAGAATA 277
 Db 246 ATTATAAGTATCATGTAAACACACACATCTTTTGTAGTTTGTACCACTACGAGTA 187

RESULT 2
 US-10-190-312A-119/c
 ; Sequence 119, Application US/10190312A
 ; Publication No. US20030199468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chromagenics B.V.
 ; APPLICANT: Otte, Arie P.
 ; APPLICANT: Kruckeberg, Arthur L.
 ; TITLE OF INVENTION: DNA sequences comprising gene transcription regulatory qualities
 ; TITLE OF INVENTION: methods for detecting and using such DNA sequences
 ; FILE REFERENCE: 2183-4993.1
 ; CURRENT APPLICATION NUMBER: US/10/190,312A
 ; CURRENT FILING DATE: 2002-07-05
 ; PRIOR APPLICATION NUMBER: 60/303,199
 ; PRIOR FILING DATE: 2001-07-05
 ; NUMBER OF SEQ ID NOS: 1079
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 119
 ; LENGTH: 1140
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: sequence of STAR A35
 US-10-190-312A-119

Query Match 7.2%; Score 64.8; DB 15; Length 1140;
 Best Local Similarity 60.0%; Pred. No. 0.0011;
 Matches 108; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 Qy 98 TAATAAAAAAGTTCACATATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGTACTAT 157
 Db 366 TCATAGCTCAATGCACCGCGGGGAATCGAACCGGGTCTGTACTATGGCAAGTACTAT 307
 Qy 158 TCTACCACTAGACCACTGGTCTTTTGTGATTAACACGACTAAATATATATATAGAT 247
 Db 306 TCTACCACTAGACCACTGGTCTTTTGTGATTAACACGACTAAATATATATATAGAT 247
 Qy 218 TTTAATTATATTTTGCACGAAATAACCGACCAAGCTTGCTGCAATTTTATTAAGAATA 277
 Db 246 ATTATAAGTATCATGTAAACACACACATCTTTTGTAGTTTGTACCACTACGAGTA 187

RESULT 3
 US-10-424-599-109423
 ; Sequence 109423, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 109423
 ; LENGTH: 2131
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_69823C.1
 US-10-424-599-109423

Query Match 7.1%; Score 64.2; DB 16; Length 2131;

Best Local Similarity 74.3%; Pred. No. 0.0018;
 Matches 81; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 Qy 98 TAATAAAAAAGTTCACATATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGTACTAT 157
 Db 1936 TTAAATTAATGCACCGCGGGATCGAACCGGGTCTGTACTATGGCAAGTACTAT 1995
 Qy 158 TCTACCACTAGACCACTGGTCTTTTGTGATTAACACGACTAAATATATATATAGAT 206
 Db 1996 TCTACCACTAGACCACTGGTCTTTTGTGATTAACACGACTAAATATATATATAGAT 2044

RESULT 4
 US-10-311-455-58/c
 ; Sequence 58, Application US/10311455
 ; Publication No. US20030143606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
 ; TITLE OF INVENTION: Cytosine methylation
 ; FILE REFERENCE: 5013.1014
 ; CURRENT APPLICATION NUMBER: US/10/311,455
 ; CURRENT FILING DATE: 2002-12-16
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537
 ; PRIOR FILING DATE: 2001-07-02
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 2424
 ; SEQ ID NO 58
 ; LENGTH: 5945
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-311-455-58

Query Match 7.1%; Score 64; DB 15; Length 5945;
 Best Local Similarity 45.7%; Pred. No. 0.0029;
 Matches 223; Conservative 0; Mismatches 265; Indels 0; Gaps 0;
 Qy 32 TCAATTAATTAATCAAAAAAATATTGTAAAAAACCACCAAGTTGATCGGTATTATTA 91
 Db 5540 TAATATACCATCAACAAAAAATCCATAAAAAATATCTAAAAAATTAATAAATAT 5481
 Qy 92 ATTATGTAATAAAGATTCACTATCTGGGAATCGAACCGGGTCTGTACTATGGCAAGA 151
 Db 5480 CAAAAATAAAAAAATCTTCTTAATAAAAAATCAATAAAAAATCAATAAAAAATTA 5421
 Qy 152 TACTATCTACCACTAGACCACTGGTCTTTTGTGATTAAGACTGTCTTTTATTGATTT 211
 Db 5420 ATTAAATAAACCTTAAAAAATAAATAATCAATTTTCAACATTAATAAAAAAATAAT 5361
 Qy 212 ATACTCTTTAATTAATTTTGCACGAAAAATACCGCAAAAGTTGGTCTGATTTATTA 271
 Db 5360 AAATATTAAAAACATTTTCAACTAACCAACCAACCAATCCCAACCAAAAAATAAAA 5301
 Qy 272 AAAGTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 331
 Db 5300 AAATCTCAAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5241
 Qy 332 CCAATTTTGGTAGTTTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 391
 Db 5240 ACTACTCTAAAAAATAATATATCTAATATTAATTAATTAATTAATTAATTAATTAATTA 5181
 Qy 392 AGTTAGTCTGGTCTTTTGAACATAAATTTCCGGGACTCAAAATAGTTTCCCGCATTT 451
 Db 5180 AATTACTATCATTTTAAATATCTAATAAATACCAAAACATTTAAAAAATTAATTTCCATTT 5121
 Qy 452 TTGCGCCAAAGAAAAACCGCAAAAGTTGGTCTGTTTCTGTAATAAAAAAATAAATAA 511

Db 2573855 AAATAAAATATAATACAAATTTAAATAATCTTAAATTTTAAATCTCTAAATATTTTAT 2573796
QY 345 GTTTTTTAAATATAATTTTATTTTAAATTTGAAATACTAAACCAAGTTAGTCGGTT 404
Db 2573795 TTTTATTTAAATATTTTAAATATATATATTTATATA-TATATATAATATATATAT 2573737
QY 405 TCTTGAAACATAATTTTCGGGACTCAAAATAGTTTCGGCATTTTTCGGCCCAAGAA 464
Db 2573736 ATACAATACAAAATAAATTTAAATAACAATTTCTATTCCTCTCCCTCCCAATAAAA 2573677
QY 465 AACCGACCAAGTTGTCGGTTTCGTAAATAAAATTTTAAATAATATATTTTAAA 524
Db 2573676 ATAAAAATCTTTACCAAAAAAATAAAAAAATAAAAAAATAAAAAAACAACA 2573617
QY 525 AA 526
Db 2573616 TA 2573615

RESULT 10
US-10-021-323-93/c
; Sequence 93, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 93
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(552)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3825-001-Q1-K6-C5
US-10-021-323-93

Query Match 6.2%; Score 56.2; DB 17; Length 552;
Best Local Similarity 44.8%; Pred. No. 0.052;
Matches 202; Conservative 0; Mismatches 249; Indels 0; Gaps 0;
QY 85 TATTTTAAATATCTAATAAAGATTCACTATCTGGGAATCGAACGGGGTCTGTACTAT 144
Db 470 TTTTITTTTTTTTTTTNAAAAATATTTTTTTTTTNNNACATAAACCCTCCAAAAA 411
QY 145 GCGAAGATATCTACCACTAGACCATGGTTCATTTTGTGTTTAAGACTGTCTTTAT 204
Db 410 ANNTTTAAATTTTTTTTTTTTAACTCCCTCCCTTTTTTTTTTCAAAAAATTTTTTT 351
QY 205 TCGATTTACTCTTTAAATATATTTTTCGACGAAATTAACGACCAAGTTGTCGAT 264
Db 350 TCTCTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTCGCCCTTTTTTTTTTTTTTTTTT 291
QY 265 TTATTAAGAGTAAATTAATCTTACCAAGTTGTCGATTTTTTTTAAATGATCGCGCAAT 324
Db 290 TAAITTTTTTTTTTTTTTTTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTCCTTTTT 231
QY 325 TAACCGACCAATTTGGTAGTTTTTTTAAATATATTTTTTTATTTTAAATGAAAA 384
Db 230 TTTTITTCACCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAA 171
QY 385 CTAAACCAAGTTAGTCGGTTCTTGAACATATAATTTTCGGGACTCAAAATAGTTCC 444

Db 170 AAAATTTATATT 111
QY 445 CGCATTTTTCGCCCAAGAAACCGACCAAGTTGTCGGTTTCGTAATAAAAAA 504
Db 110 ATTTTTTTTTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 51
QY 505 TTTAAAAATATATTTTAAAAAACCGACCA 535
Db 50 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 20

RESULT 11
US-10-311-455-1749/c
; Sequence 1749, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1749
; LENGTH: 6239
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1749

Query Match 6.1%; Score 55.4; DB 15; Length 6239;
Best Local Similarity 46.9%; Pred. No. 0.19; Indels 0; Gaps 0;
Matches 173; Conservative 0; Mismatches 196;
QY 172 ATGGTTCATTTTGTTTTAAGACTGTCTTTTATTTGATTTTATCTCTTTTATATATTTT 231
Db 1844 ATATATTTTTTTTATTTTAAATTAACGTTTATATATTTTATATTTTAAATTTCTT 1785
QY 232 TGCACGAAATPAACCGACCAAGTTGTCGATTTTATTAAGATAAAATTAATACTACTACCA 291
Db 1784 TTATTAATATAATAAATAATACTTATTTATTTATATATATATATATATATATATA 1725
QY 292 AGTTGGTCGATTTTTTTTAAATGATCCGCGAATTAACGACCAATTTGGTAGGTTTTTT 351
Db 1724 TCTTATATTTTTTTTTTAAATTTAAATTTATTTTATTTTATTTTATTTATTTATTT 1665
QY 352 TAAATTTAAATTTTATTTTATTTTAAATTTGAAAACTAACCAAGTTAGTCGGTTTCTTCAA 411
Db 1664 TATAATTAACAATAATCATCATACACCTATAATCCCACTACAAAAAATTTACTTAAAC 1605
QY 412 ACATAAATTCGGGACTCAAAATAGTTTCCGCAATTTTTCGCCCAAGAAAAACCGAC 471
Db 1604 CCAAAAAATCGAAACTCAACAAACCAATAATATATCACTATATCTCAACCTCAACGACA 1545
QY 472 CAAGTTGGTTCGGTTTCGTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 531
Db 1544 AATAAACCCTACTCTCTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1485
QY 532 CCAACTTTA 540
Db 1484 TTAICTTTA 1476

RESULT 12

	Query Match	6.1%	Score 55.2	DB 15	Length 5407
	Best Local Similarity	47.6%	Pred. No. 0.2	Indels	0
	Matches 162	Conservative 0	Mismatches 178	Indels	Gaps 0
Qy	186	TTTTAAGACTGCTCTTTTAAATTTGATTTATATCTCTTTAAATTTATATATTTTGCAGAAATTAAC	245		
Db	2083	TATTTATAATATCTCTATACCCACCACTCTTTTAAATATTTTTTAAATAAAAAAAC	2024		
Qy	246	CGACGCAAAAGTTGGTCGATTTTATTAATAAAGTAAAAATTACTTACCAAGTTGGTCGATTTT	305		
Db	2023	ATTTCGAAATTTAAATTTATACAAATTCGAAATTTAAATCTTCAAAAACTCACTTATTTT	1964		

306	TTTAAATGATCGCGGAAATTAACCGACCAATTTTGGTAGGTTTTTTTAAATAATTAATTTTT	365
Qy		
1963	AATAAAAAAATAACCCAAATATATCTACCACTTTTCCTAAACTCTATTTTAAANAATTTCTAC	1904
Db		
366	ATTTATTTTAAATGAAAAAAGCTAACCAAAAGTTAGTCGCTTTCTTGAACAATAAAATTCGCG	425
Qy		
1903	CCAAATAATAAAAAAGTATAACACAAAAATAATAATTTTCATATCTATAACCAACCAATAT	1844
Db		
426	GGACTCAAAAAATAGTTTTCCCGGCATTTTTTGGCCCAAGAAAAACCGACCAAGTTGGTCGGT	485
Qy		
1843	TAAATCTCTACTATTTATCTATACTATTTATCCAAAAAATAAATTCAAATTCCTATCATATAT	1784
Db		
486	TTCTGAAAAAATAAAAAAATTTTAAAAAATAATTTTTAAAA	525
Qy		
1783	TAAAAATCTAAAAATATTTTAAAAAATAATATACATTAATTTAAAA	1744
Db		

RESULT 14
US-09-834-975-47/c
Sequence 47, Application US/09834975
Patent No. US20020110815a1
GENERAL INFORMATION:
APPLICANT: Lallie, James
APPLICANT: Brown, Jeffrey
APPLICANT: Bolt, Andrew
APPLICANT: Van Huffel, Christophe
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
FILE REFERENCE: MRI-016B
CURRENT APPLICATION NUMBER: US/09/834,975
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/197,538
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 1046
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 47
LENGTH: 368
TYPE: DNA
ORGANISM: Homo sapiens

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 06:56:19 ; Search time 4307.14 Seconds
(without alignments)
10990.348 Million cell updates/sec

Title: us-10-009-966c-2_COPY_2155_3155
Perfect score: 1001
Sequence: 1 aacatagacttttgattaaa.....accactactagaatccgg 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.btg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000.2	99.9	4312	6	AX068013 Sequence
2	997	99.6	3294	6	AX068012 Sequence
3	997	99.6	4135	6	AX068019 Sequence
4	980.6	98.0	3293	8	AF376772S1 Nicotiana
5	100	10.0	3663	8	AB012639 Nicotiana
6	66.6	6.7	154256	2	CR678065 Danio rer
7	65.6	6.6	79018	9	AL627107 Human DNA
8	64.6	6.5	207772	2	EX897691 Danio rer
9	64.2	6.4	201329	2	CR450778 Danio rer
10	64	6.4	156463	2	CR388040 Danio rer
11	64	6.4	160534	2	CR450837 Danio rer
12	63.6	6.4	201312	2	AC067900 Homo sapi
13	63	6.3	154675	2	CR354440 Danio rer
14	62.8	6.3	4753	8	NS1252274 Nicotiana
15	62.6	6.3	90550	9	AL592166 Human DNA
16	62.6	6.3	174001	2	EX927081 Danio rer
17	62.6	6.3	193764	2	CR450831 Danio rer
18	62.4	6.2	172811	9	AC066608 Homo sapi
19	62.4	6.2	236582	2	CR388102 Danio rer

20	62.4	6.2	349418	3	CR382398	Plasmodiu
21	61.8	6.2	185672	2	BX640474	Danio rer
22	61.4	6.1	177888	2	EX957314	Danio rer
23	61.2	6.1	245072	2	CR381551	Danio rer
24	61.2	6.1	348034	3	CR382400	Plasmodiu
25	61	6.1	163533	3	CR627488	Danio rer
26	61	6.1	190128	2	BX908801	Danio rer
27	61	6.1	233658	2	CR354421	Danio rer
28	60.8	6.1	186313	2	CR631129	Danio rer
29	60.6	6.1	181539	2	EX957351	Danio rer
30	60.4	6.0	1141	6	AX083744	Sequence
31	60.4	6.0	125836	2	AC137629	Homo sapi
32	60.4	6.0	156978	2	CR339045	Danio rer
33	60.2	6.0	110000	3	PFMAL1P2_1	Continuation (2 of
34	60.2	6.0	132112	2	CR628662	Danio rer
35	60	6.0	138982	2	CR388017	Danio rer
36	60	6.0	196756	2	CR556696	Danio rer
37	59.8	6.0	72243	9	AL731858	Human DNA
38	59.8	6.0	155156	2	CR627499	Danio rer
39	59.8	6.0	162206	9	AC108706	Homo sapi
40	59.8	6.0	177423	2	CR356242	Danio rer
41	59.8	6.0	250663	3	AE014826	Plasmodiu
42	59.8	6.0	313050	3	PF929352	Plasmodiu
43	59.6	6.0	189963	2	CR559944	Danio rer
44	59.4	5.9	198340	2	CR556709	Danio rer
45	59.2	5.9	148856	2	BX942828	Danio rer

ALIGNMENTS

RESULT 1
AX068013
LOCUS AX068013 4312 bp DNA linear PAT 19-JAN-2001
DEFINITION Sequence 2 from Patent WO0077187.
ACCESSION AX068013
VERSION AX068013.1 GI:12329821
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum

REFERENCE 1
AUTHORS Roitsch, T.D.
TITLE Promoter system and production and use of the same
JOURNAL Patent: WO 0077187-A 2 21-DEC-2000.
Roitsch, Thomas, Dr. (1997)

FEATURES
source
1. .4312
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"

ORIGIN

Query Match	99.9%;	Score	1000.2;	DB	6;	Length	4312;	
Best Local Similarity	100.0%;	Pred. No.	7.6e-172;					
Matches 1001;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	AATATAGACTTTTGATTAAATTAATTAATTTGTATGACAAAGGATTAAACCTAGTTAAT	60					
Db	2155	AATATAGACTTTTGATTAAATTAATTAATTTGTATGACAAAGGATTAAACCTAGTTAAT	2214					
Qy	61	GATAATAACTTCAGATCTATATCAATTAATTAACAGCAATCACCGTCTATACAGCGTTGAGAG	120					
Db	2215	GATAATAACTTCAGATCTATATCAATTAATTAACAGCAATCACCGTCTATACAGCGTTGAGAG	2274					
Qy	121	AGATTAAATGTCATGTCATTCATTAATTTCAAGATCAATTAATGATAGGGGAATATCAAG	180					
Db	2275	AAGATTAAATGTCATGTCATTCATTAATTTCAAGATCAATTAATGATAGGGGAATATCAAG	2334					
Qy	181	CAATAAATAACGATAAATGGCATTAATAAGTAAATTAAGGAGAGATGATTCACCCAAATTGAA	240					

in start get
parent

8 ACTTCAGATCTATAATCAATTAAACAGCAATCACGGTCATAGCAGCGTTGAG


```

/note="assembly_fragment:00168
fragment_chain:2"
misc_feature
146937..1511721
/note="assembly_fragment:00105.0"
misc_feature
151822..156463
/note="assembly_fragment:00049
clone_end:T7
vector_side:right"
ORIGIN
Query Match 6.4%; Score 64; DB 2; Length 156463;
Best Local Similarity 47.2%; Pred. No. 0.021;
Matches 222; Conservative 1; Mismatches 246; Indels 1; Gaps 1;
Qy 459 ATTATATGGGACATCTCCTCAATCAATCCATAAGTACATACACCAAGATATATCAATAA 518
Db 112418 ATTCAGCTTGATTTTCATGGCCAAATGTATGAAACTTGTATCAAAAAAATGATACA 112359
Qy 519 AATATTTTTTGAATATCTTATTATAAAACTAGCTGTTAGCACTCGACCTGGTCGYTA 578
Db 112358 AAAATTTGATACAAAATGATCAAAATATACGGTATTTGATCATGCGATCTATTACAA 112299
Qy 579 TTGACTACTCGTTACGAGCCCTGCTCAATTAATCGACCTCGATTACATCACTTTCTTA 638
Db 112298 GATATCCTATTTATTCATAAGGTTCAATTTATTTGTTTCATCGGTCAGGCTACAGTGAT 112239
Qy 639 CGATCTGCTTCATGTCATCAATCTTAATGAAGACAGATTTTGACCCATACAAATATGAC 698
Db 112238 TAGTAGGCTTTAATTTATAGCAAGTAGAAGAG-TTTCGCTCATAAACCAATCCCC 112180
Qy 699 AAAATTTGCTTCCAAAGAAAACATGGCTCTTATAGTGAATATGCTTGAGCTGTTATAGAA 758
Db 112179 AAACAGCTGTCAAAAAABAAAATATATATATATATATATATATATATATATATAT 112120
Qy 759 AGATCTGAATTTATTTATAGATAGATGTTTTTTTCTTTCTTTTCTTTTCTTTTCTTTTCT 818
Db 112119 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 112060
Qy 819 AAGCAACCATGAATAGAAAAGGCTTACTATATATATCAAGGAATGGTGTGTTTTCTT 878
Db 112059 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 112000
Qy 879 TAAATATGGATAAAATTTTGATATATAGAAATAGATCAATTAACAA 928
Db 111999 TATATATAAATAAAAAATATATATATATATATATATATATATATATATATATATA 111950

RESULT 11
CR450837 160534 bp DNA linear HTG 24-MAY-2004
LOCUS Danio rerio clone CH211-51L13, *** SEQUENCING IN PROGRESS ***, 10
DEFINITION unordered pieces.
ACCESSION CR450837
VERSION CR450837.1 GI:47604495
KEYWORDS HTG; HTGS; PHASEL.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 160534)
Sims,S.
REFERENCE Direct Submission
AUTHORS Submitted (21-MAY-2004) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
COMMENT zfish-help@sanger.ac.uk UK clone requests: clonesrequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zC51113
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 156811 bases at least Q40
Consensus quality: 157308 bases at least Q30
Consensus quality: 157864 bases at least Q20
Insert size: 159634; sum-of-contigs
Insert size: 176376; 2.7% error; agarose-fp
Quality coverage: 8.20x in Q20 bases; sum-of-contigs Quality
coverage: 7.72x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 8994: contig of 8994 bp in length
* 8995 3094: gap of 100 bp
* 9095 31712: contig of 22818 bp in length
* 31713 31812: gap of 100 bp
* 31813 50359: contig of 18547 bp in length
* 50360 50459: gap of 100 bp
* 50460 56578: contig of 6119 bp in length
* 56579 56678: gap of 100 bp
* 56679 61867: contig of 5188 bp in length
* 61867 89747: contig of 27780 bp in length
* 89747 130975: contig of 41129 bp in length
* 130976 131075: gap of 100 bp
* 131076 143529: contig of 12454 bp in length
* 143530 143629: gap of 100 bp
* 143630 148065: contig of 4435 bp in length
* 148065 148164: gap of 100 bp
* 148165 160534: contig of 12370 bp in length.
FEATURES
Location/Qualifiers
1..160534
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-51L13"
/clone_lib="CHORI-211"
1..8994
/note="assembly_fragment:00306
fragment_chain:1"
misc_feature
9095..31712
/note="assembly_fragment:01004
fragment_chain:1"
misc_feature
31813..50359
/note="assembly_fragment:01306
fragment_chain:1
clone_end:T7
vector_side:left"
misc_feature
50460..56578
/note="assembly_fragment:00219
fragment_chain:1"
misc_feature
56679..61866
/note="assembly_fragment:00108
fragment_chain:2"
misc_feature
61967..89746
/note="assembly_fragment:00715
fragment_chain:2"
misc_feature
89847..130975
/note="assembly_fragment:01645
fragment_chain:3"
misc_feature
131076..143529
/note="assembly_fragment:00557
fragment_chain:3"
misc_feature
143630..148064
/note="assembly_fragment:00162.0"

```

```

misc_feature 148165..160534
/notes="assembly_fragment:00424"
ORIGIN
Query Match 6.4%; Score 64; DB 2; Length 160534;
Best Local Similarity 34.7%; Pred. No. 0.021;
Matches 215; Conservative 0; Mismatches 401; Indels 3; Gaps 1;

QY 313 ATCTAATGAAAAAGTATGGAATAGTAGAATATCGAATCTCTTTGAAAAGGTAGTGATG 372
Db ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 9997
QY 373 TCTTTATCTAGAGAGAAAGTCGCTTTCCTCAAGAAATATTTTATCAGAGAAATATACAT 432
Db NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9057
QY 433 CCCCTCTCTCCCTATCTCTTTTCTATTTATATGCGACATTCCTCAATCAATCTCTAA 492
Db NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9117
QY 493 GTACATACACCAAGAAATATTCATATAAAATATTTTGTGAATATCTTATATATAAACTAG 552
Db TTAATATTAATTAATTAATTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9177
QY 553 CTGTT--ACCACTCGACCTCGCTCGTATGACTACTCGGTTACGAGCCCTGCTCATTTA 609
Db TTTTGTAAATAATATAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9237
QY 610 CTAACTGACCTGATTCATCACTCTCTACGACTGCTCTGCTCAATCTCTTAATGAAA 669
Db TTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9297
QY 670 GCAGATTTGACCAATATATATGACAAATTCCTTCCAAAGAAACATGGCTCTTA 729
Db NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9357
QY 730 TAGTGAATATCGTTAGCTGTTATAGAAGATCTGAATTTATTTTATAAGATAGTGTT 789
Db TATATATATATATATATATATATATATATATATATATATATATATATATATATATA 9417
QY 790 TTTTCTTTCTTTTCATCTAAGAGTAAAGCAACCATGAATAGAAAAGGCTTAGTAA 849
Db TATATTGTTNAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 9477
QY 850 TATATATCAAGAAAGGCTTTTCTTTTAAATATGATGATAAAATTTGTGAATAGAA 909
Db TTAATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATAT 9537
QY 910 GATTAGATCAATTAACAAA 928
Db ATATATATATATATATATA 9556

RESULT 12
AC067900 201312 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 4 clone RP11-775M3 map 4, WORKING DRAFT
DEFINITION SEQUENCE, 5 unordered pieces.
ACCESSION AC067900
VERSION AC067900.3 GI:12545341
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 201312)
Birken, B., Linton, L., Nussbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 4, clone RP11-775M3
JOURNAL Unpublished
REFERENCE 2. (bases 1 to 201312)
Birken, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Bouckghalter, B., Brown, A., Burkett, G.,

```

```

Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galligan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Liecu, C., Liu, G., Locke, K., Macdonald, P., Marquis, J.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meidrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Navlor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, E., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 201312)
Birken, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Boguslavsky, L., Bouckghalter, B., Brown, A.,
Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Fato, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Lehoczy, J., Karatas, A., LaRocque, K., Lamazares, R., Landers, T.,
Lehoczy, C., Levine, R., Liu, G., MacLean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meidrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, J., Navlor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Riack, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 26, 2001 this sequence version replaced gi:8138718.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIER
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7524
Center clone name: 775 M.3
----- Summary Statistics
Sequencing vector: M13; M7815; 41% of reads
Sequencing vector: Plasmid; n/a; 59% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 198330 bases at least Q40
Consensus quality: 199387 bases at least Q30
Consensus quality: 200109 bases at least Q20
Insert size: 21000; agarose-gel
Insert size: 200912; sum-of-contigs
Quality coverage: 9.1 in Q20 bases; agarose-gel
Quality coverage: 9.5 in Q20 ba.

```

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 77871: contig of 77871 bp in length
 77872 77971: gap of 100 bp
 77972 83159: contig of 5188 bp in length
 83160 83259: gap of 100 bp
 83260 115460: contig of 32201 bp in length
 115461 115560: gap of 100 bp
 115561 115895: contig of 43435 bp in length
 115896 159095: gap of 100 bp
 159096 201312: contig of 42217 bp in length.

FEATURES

source

1. 201312
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="4"
 /map="4"
 /clone="RP11-775M3"
 /clone_lib="RPC1-11 Human Male BAC"

misc_feature

1. 77871
 /note="assembly_fragment
 clone end:SP6
 vector side:left"
 77972..83159
 /note="assembly_fragment"
 83260..115460
 /note="assembly_fragment"
 115561..115895
 /note="assembly_fragment"
 159096..201312
 /note="assembly_fragment"

ORIGIN

Query Match 6.4%; Score 63.6; DB 2; Length 201312;
 Best Local Similarity 39.8%; Pred. No. 0.023;
 Matches 371; Conservative 1; Mismatches 555; Indels 5; Gaps 3;

QY 2 ATATAGACTTTTGATTAAATTAATATATATGATGACAAAGGATTAAACCTAGTTAAG 61
 DB 158594 ATATATATAAATATATATAAATATATATAAATAAATATATATAAATATATA 158653
 QY 62 ATAATAACTTCAGATCTATCAATTAACAGCAATCACGGTCATAGCAGCGTTGAGAGA 121
 DB 158654 AATATATAAACAATATAAACAATATATAAATATATAAATATATAAATATATA 158713
 QY 122 AGATTAATGTGATGTCATCAATATTTCAAGATCAATTAATGATAGGGGAATCAAGC 181
 DB 158714 TAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATAA 158773
 QY 182 AATAAATAACGATAAATGGGATTAAGCTAAATTAAGGAGATGATTCACCAATATTGAAT 241
 DB 158774 AATATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 158833
 QY 242 GAGTGGATGATCTCTTTTGGCAATGATGATGCGGCAATGCTAGAGATGTTGGG 301
 DB 158834 AATATGATAAATATATAAATATATAAATATATAAATATATAAATATATAA 158893
 QY 302 ACCCTTCGGATCTATGAAAAGATGGAATAGTATCGAATCTCTTTAGAA 361
 DB 158894 TATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 158953
 QY 362 GGTAGTGATGCTCTTTTATCTAGAGAGAAA-GTCTGCTTTTCAAGAATATTTTATCAG 420
 DB 158954 AAAGTATATAAATATATAAATATAGATAAATAAATAAATAAATAAATAAATA 159013
 QY 421 AGATATTACATCCCTCTCTCCCTATCTCTTTTCTATTTATATGGGACATTCCTCAA 480
 DB 159014 NNN 159073

QY 481 TCAATCCTAAAAGTACATACACCAAGATATTCAATAAAATATTTTGTGAATTTCTAT 540
 DB 159074 NNN 159133
 QY 541 TATAAAAAGTACTAGCTGTTAGCACTCGACCTCGGTCGTTATTTGACTACTCGGTTACGAGCCC 600
 DB 159134 TATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 159193
 QY 601 TGTCAATTTACTAATCGACCTCGATTACATCACTTCTACGACTACTGCTTCATGTCAAATC 660
 DB 159194 TATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 159253
 QY 661 TTAATGAAAGCAGAGATTTTGACCCATACAAATAATATGACAAAATTCCTTCCAAAGAAACA 720
 DB 159254 TATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 159313
 QY 721 TGGCTCT-TATAGTGAATATCGTTAGACTGTTTATAGAAAGATCTGAAATTTATTAAG 779
 DB 159314 TAAATATATATAAATATATAAATATATAAATATATAAATATATAAATATATAA 159373
 QY 780 AATAGTGTTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 839
 DB 159374 TATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 159433
 QY 840 GCTTAGTAACATATATAT---CAAAGGAATGGTGTCTTTTCTTAAATATGGAATAAATT 896
 DB 159434 TATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 159493
 QY 897 TGTGAATATAGAAGATTAGATCAATTAACAA 928
 DB 159494 TATATATAAAGATGTTGTTATTTTAAAA 159525

RESULT 13

CS354440/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

CR354440 154675 bp DNA linear HTG 27-MAR-2004
 Danio rerio clone DKEYP-100B9, *** SEQUENCING IN PROGRESS ***, 13
 unordered pieces.

CR354440
 CR354440.6 GI:45772251
 HTG; HTGS PHASE1.

Danio rerio (zebrafish)
 Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 154675)

McLay, K.
 Direct Submission

Submitted (28-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

fish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Mar 27, 2004 this sequence version replaced gi:45723965.

----- Genome Center
 Center: Wellcome Trust Sanger Institute

Center code: SC
 Web site: http://www.sanger.ac.uk

Contact: zfish-help@sanger.ac.uk
 ----- Project Information

Center project name: zkp100B9
 ----- Summary Statistics

Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads

Consensus quality: 149604 bases at least Q40
 Consensus quality: 150352 bases at least Q30

Consensus quality: 151128 bases at least Q20
 Insert size: 153475; sum-of-contigs

Insert size: 163247; 2.2% error; agarose-
 Quality coverage: 8.23x in Q20 bases; sum-of-contigs Quality

coverage: 7.94x in Q20 bases; agarose-

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 13 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2783: contig of 2783 bp in length
2784 2883: gap of 100 bp
2884 13276: contig of 10393 bp in length
13277 13376: gap of 100 bp
13377 15922: contig of 2546 bp in length
15923 16022: gap of 100 bp
16023 25315: contig of 9293 bp in length
25316 25415: gap of 100 bp
25416 34278: contig of 8863 bp in length
34279 34378: gap of 100 bp
34379 38937: contig of 4559 bp in length
38938 39037: gap of 100 bp
39038 47656: contig of 8619 bp in length
47657 47757: gap of 100 bp
47758 57042: contig of 9286 bp in length
57043 57142: gap of 100 bp
57143 64811: contig of 7669 bp in length
64812 64911: gap of 100 bp
64912 105347: contig of 40436 bp in length
105348 105447: gap of 100 bp
105448 111960: contig of 6513 bp in length
111961 112060: gap of 100 bp
112061 150521: contig of 38461 bp in length
150522 150621: gap of 100 bp
150622 154675: contig of 4054 bp in length.

FEATURES

source

1..154675
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEYP-100B9"
/clone_lib="DanioKeyPilot"

misc_feature

1..2783
/note="assembly_fragment:00085
fragment_chain:1"

misc_feature

2884..13276
/note="assembly_fragment:00936
fragment_chain:1"

misc_feature

13377..15922
/note="assembly_fragment:00113
fragment_chain:1"

misc_feature

16023..25315
/note="assembly_fragment:00786
fragment_chain:1"

misc_feature

25416..34278
/note="assembly_fragment:00646
fragment_chain:1"

misc_feature

34379..38937
/note="assembly_fragment:00260
fragment_chain:1"

misc_feature

39038..47656
/note="assembly_fragment:00423
fragment_chain:1"

misc_feature

47757..57042
/note="assembly_fragment:00533
fragment_chain:1"

misc_feature

57143..64811
/note="assembly_fragment:00325
fragment_chain:1"

misc_feature

64912..105347
/note="assembly_fragment:01784
fragment_chain:1"

misc_feature

105448..111960
/note="assembly_fragment:00198
fragment_chain:1"

misc_feature

112061..150521
/note="assembly_fragment:01096

fragment_chain:1"
150622..154675
/note="assembly_fragment:00152
fragment_chain:1"

ORIGIN

Query Match 6.3%; Score 63; DB 2; Length 154675;
Best Local Similarity 39.0%; Pred. No. 0.032;
Matches 218; Conservative 0; Mismatches 340; Indels 1; Gaps 1;
QY 313 ATCTAATGAAAAAGTATGGAATAGTAGAATATCGAATCTCTTTAGAAAGGTAGTGATG 372
DB 34615 ATATATATATATATATAAAATATATTAATAATATATATATATATATATATATAT 34556
QY 373 TCTTTTATCTAGAGAGAAGTCGCTTTTCTAAGAAATATTTTATCAGAGAAATATACAT 432
DB 34555 ATTTATATATTTATTAATATATATATTTATATATATTTAATAAAAAA 34496
QY 433 CCCCTCTCTCCCTATCTCTTTTCTATTATATGCGACATTCCTCAATCCTAAAA 492
DB 34495 AATAATATATTTATAAATATATATATTAATAATTTTATAATTTTATAAATTTTATA 34436
QY 493 GTACATACACCAAGAAATATTCATAAAATATTTTTTTCG-AAATATCTATTATAAAACTA 551
DB 34435 TTATTTTAAAAAATATAATATATAAATAATAATTTATAATAATTTATATATTTNN 34376
QY 552 GCTGTAGCACTCGACCTCGTCGTYATTGACTACTCGTTACGAGCCCTGCATTTACT 611
DB 34375 NNN 34316
QY 612 AATCGACCTCGATTACATCACCTTCTAGGATACTGCTTCAATCTTAATGAAAGC 671
DB 34315 NNN 34256
QY 672 AGATTTTGACCATACAATAATATGACAAATGCTTCCAAAGAAAAAATGCTCTTATA 731
DB 34255 TATTTATATTTATTTTATATATATATATATTTTATTTATTTATATATTTTATA 34196
QY 732 GTGAAATATCGTTAGACTGTTATAGAAAGATCTGAATTTTATTAAGAATAAGTCTTT 791
DB 34195 TA 34136
QY 792 TTTCTTTCTTTTCATATCTAAGGAGTAAAGCAACCATGAATAGAAAGCTTTAGTA 851
DB 34135 TA 34076
QY 852 TATATCAAGGAATCGTGT 870
DB 34075 TATATAAAGTTTAAAGT 34057

RESULT 14

NSY292274/c 4753 bp DNA linear PLN 12-JAN-2001
LOCUS Nicotiana sylvestris DNA flanking T-DNA insert line S22KdotNOSpro
DEFINITION orphan R.
ACCESSION AJ292274
VERSION AJ292274.1 GI:12227269
KEYWORDS
SOURCE Nicotiana sylvestris (wood tobacco)
ORGANISM
REFERENCE 1
AUTHORS Kunz, C., Naragavavana, J., Jakowitsch, J., Park, Y. D., Delon, T. R., Kovarik, A., Koukalova, B., van der Winder, J., Mette, M. F., Aufeatz, W., Matzke, M. A. and Matzke, A. J.
TITLE Contaminating bacterial transposon sequences as possible contributors to silencing and methylation of single copy transgenes in plants
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4753)

AUTHORS Winden and van der, J.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-2001) Winden, van der J., Plant Molecular Genetics, Institute Of Molecular Biology, Billrothstrasse 11, A-5020 Salzburg, AUSTRIA

FEATURES
source Location/Qualifiers
1..4753
/organism="Nicotiana sylvestris"
/mol_type="genomic DNA"
/db_xref="taxon:4096"
misc_feature 1..4753
/note="DNA flanking T-DNA insert line S22KdotNOSpro orphan R"

ORIGIN
Query Match 6.3%; Score 62.8; DB 8; Length 4753;
Best Local Similarity 52.7%; Pred. No. 0.072;
Matches 277; Conservative 1; Mismatches 233; Indels 15; Gaps 6;
QY 5 TAGACATTTTCATTAAATTAATTAATATTTGATGACAAAGGATTAAACCTAGTTAAATGATA 64
DB 2298 TAAGCTTTATGTTAACTAATTAATGAGGA--ATAGAGGGAGAACTCTAGTTAAACATA 2241
QY 65 ATAACCTTCAGATCTATATAATCAATTAACAGCAATCACGTTCATAGCAGCGTTGAGAGAAGA 124
DB 2240 ATAAATCCTAGATAAAAAAAGCTGGGGTATGTGAGGATATAAGACCGGTTCATATGATGAT 2181
QY 125 TTAAGTGTGATGYCATTCATATTTCAAGATCATTAATGATAGGGGAATATCAAGCAAT 184
DB 2180 AAATATCGAATATGTGTTCATGATGTCATGAATGAATAAGTAA---TAGAAGCATGCAAT 2124
QY 185 AAATAACGATAAATGGCATTAAAGTAATAAAGAGAGAAATGATTCACCC--AATATTGAATGA 243
DB 2123 GAATGTAGAGATGACAAATAAATGTAATAAGAGAGAAATTTATCACCACAAATATTAGGTGA 2064
QY 244 GGTGGATGATCTTCCTTTT--TGCAATGATGAATGATGGGCAATACTAGAAATGTTGGGA 302
DB 2063 CTTCGGATGATTCCTCTTCCTTCGACAAACGCTGGGATGATAGATAATGAAGATAGATAAG 2004
QY 303 CCCTCTCGGATCTAATGAAAAAAGTATGGAATAGTAGATAATCGAATCTCTTTAGAAAG 362
DB 2003 GAATCTTTGGATCTGTGA-----ATGGCGAGCTATATTGAAAAATGTAATCAGTGAA 1951
QY 363 GTAGTGATGTCCTTTTATCTAGAGAGAAAGTCGCTTTTCAAAGAATATTTTATCAGAG 422
DB 1950 CAAGACACACTTTTATATATCTCTATATTCAACCCCTTTACAATGTCCTCTATCAAAA 1891
QY 423 AATATTACATCCCCCTCTCCCTATCTCTTTTCTATTTATATGGGACATTTCTCTCAATC 482
DB 1890 AGTGA--ATCCCTCTTTTGTCTTTATCTCTCTCCCTATTTATAGGGGATATCCTTGAA 1832
QY 483 AATCCCTAAAAAGTACATACACAAAGATATTCAATAAAATATTTTT 528
DB 1831 ACCCTAAAAAGTACACTGAACAAATATTCAATGAAATATTCCT 1786

RESULT 15
AL592166 90550 bp DNA linear PRI 23-MAR-2003
LOCUS Human DNA sequence from clone Rp11-269F19 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL592166
VERSION AL592166.16 GI:29165241
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 90550)
Wallis, J.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (23-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquary@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Mar 23, 2003 this sequence version replaced gi:18476114.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
 RP11-269F19 is from the library RPC1-11.1 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBCae3.6
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquary@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion the clone being a YAC.

```

FEATURES
source
Location/Qualifiers
1..90550
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-289F19"
/clone_lib="RPC1-11.1"

```

ORIGIN	Query Match	6.3%;	Score 62.6;	DB 9;	Length 90550;
	Best Local Similarity	43.1%;	Pred. No. 0.042;		
	Matches 406;	Conservative	1;	Mismatches 525;	Indels 9; Gaps 2;
QY	2	ATATAGACCTTTTCATTAAATTAATTAATTTGATGACAAAGGATTAAACCTAGTTAATG	61		
Db	59484	ATATACATATATAAATAAATAAATAAAGTATTTTGAATATATATGATATATAAATAT	5954		
QY	62	ATAATAACTTCAGATCTATAATCAATTAACAGACCAATCACGTCCATGACGCGTTGAGAGA	121		
Db	59544	ATAAAAAATATATATAAATATATAAATATATAAATATAAATAATATAAATAAATAA	5960		
QY	122	AGATTAAATGTGANGYCATCTCAATATTTTCAAGATCATTTAATGATGCGGGAATATCAAGC	181		
Db	59604	AATATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATAAATA	5966		
QY	182	AATAAATAACGATAAATGGCATTTAAAGTAAATAAGGAGATGATTCACCCCAATTTGCAAT	241		
Db	59664	TATATAATATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAATATA	5972		
QY	242	GAGGTGGATGATCTCTTCTTTTCACAAATGATGAATGGCGCAAACTACTAGAATGTTGGG	301		
Db	59724	AATATATAAATATATAAATATAAATATAAATATAAATATAAATATAAATAAATAAATA	5978		
QY	302	ACCCCTTCTCGGATCTAATGAAAAAGATGGAATAGTAGATAAATCGAATCTCTTTAGAAA	361		
Db	59784	TATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAAATA	5984		

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 06:24:54 ; Search time 493.658 Seconds
(without alignments)
10644.348 Million cell updates/sec

Title: US-10-009-966c-2_COPY_2155_3155

Perfect score: 1001

Sequence: 1 aatagagcttttgattaa.....accactactagaatccgg 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1000.2	99.9	4312	4 AAF25736	Aaf25736 Tobacco i
2	997	99.6	3294	4 AAF25735	Aaf25735 Tobacco i
3	997	99.6	4135	4 AAF25742	Aaf25742 Tobacco p
4	55.6	5.6	158001	12 ADL17884	Adl17884 Human pho
C 5	55.2	5.5	8056	8 ABZ10246	Abz10246 Haematopo
C 6	54.8	5.5	37515	6 ABQ66998	Abq66998 Human ang
C 7	54.4	5.4	113515	6 ABL34175	Ab134175 Human imm
8	52.6	5.3	8201	6 ABL34511	Ab134511 Human met
9	52.6	5.3	8201	6 ABL70540	Ab170540 Chemicall
C 10	52	5.2	8056	8 ABZ10100	Abz10100 Haematopo
C 11	52	5.2	13584	6 ABL32615	Ab132615 Human imm
C 12	50.8	5.1	8056	8 ABZ10246	Abz10246 Haematopo
C 13	50.6	5.1	115218	8 ACA64845	Ac64845 Human HNR
C 14	50	5.0	5413	4 AAS46694	Aas46694 Tumour su
C 15	50	5.0	11745	6 ABK28332	Abk28332 DNA trans
16	49.4	4.9	11694	4 AAS46698	Aas46698 Tumour su
17	49.2	4.9	5360	6 ABL33937	Ab133937 Human imm
18	49.2	4.9	5360	6 ABL92297	Ab192297 Chemicall
19	49.2	4.9	5360	6 ABL49366	Ab149366 Human pol
C 20	49.2	4.9	5743	6 ABL34068	Ab134068 Human imm
C 21	49.2	4.9	7306	6 ABL33636	Ab133636 Human imm

C 22	49.2	4.9	34688	6 ABQ67060	Abq67060 Human ang
23	49	4.9	9810	6 ABL32427	Ab132427 Human imm
24	48.8	4.9	6109	6 ABL32326	Ab132326 Human imm
25	48.6	4.9	6109	6 AAS61077	Aas61077 Human gen
C 26	48.6	4.9	50000	6 ABL56201	Ab156201 AmEPV gen
27	48.4	4.8	12507	6 ABL32298	Ab132298 Human imm
28	48.4	4.8	20420	4 AAK73165	Aak73165 Human imm
C 29	48.4	4.8	20420	6 ABK69933	Abk69933 Human sec
30	48.4	4.8	37515	6 ABQ66997	Abq66997 Human ang
31	48.2	4.8	7461	6 ABL33784	Ab133784 Human imm
32	47.8	4.8	15732	4 AAS45388	Aas45388 Chemicall
33	47.8	4.8	15732	6 ABK28333	Abk28333 DNA trans
34	47.8	4.8	115218	8 ACA64845	Ac64845 Human HNR
35	47.2	4.7	8056	8 ABZ10100	Abz10100 Haematopo
36	47.2	4.7	19965	4 AAK73166	Aak73166 Human imm
37	47.2	4.7	19965	6 ABK69932	Abk69932 Human sec
38	47	4.7	5572	6 ABL33424	Ab133424 Human imm
39	47	4.7	5572	6 ABL54359	Ab154359 Chemicall
40	47	4.7	5979	4 AAS45313	Aas45313 Chemicall
41	47	4.7	5979	6 ABK28152	Abk28152 DNA trans
42	47	4.7	11422	6 ABK39937	Abk39937 Human che
43	47	4.7	11422	6 ABL32219	Ab132219 Human imm
44	46.8	4.7	8047	6 ABN80324	Abn80324 Human che
45	46.8	4.7	14316	6 ABK31519	Abk31519 Signal tr

ALIGNMENTS

RESULT 1
AAF25736
ID AAF25736 standard; DNA; 4312 BP.
XX
AC AAF25736;
XX
DT 06-APR-2001 (first entry)
XX
DE Tobacco invertase promoter SEQ ID NO 2.
XX
KW Tobacco; tapetum-specific; pollen-specific; promoter; seed; invertase;
KW male-sterile plant; in vitro embryogenesis; seedless fruit;
KW transgenic plant; ds.
XX
OS Nicotiana tabacum.
XX
PN WC2000077187-A2.
XX
PD 21-DEC-2000.
XX
PF 13-JUN-2000; 2000WO-DE001944.
XX
PR 12-JUN-1999; 99DE-02009998.
PR 04-APR-2000; 2000DE-02005992.
PR 26-APR-2000; 2000DE-02007494.
(ROIT/) ROITSCH T.
PI Roitsch T;
XX
XX WPI; 2001-080685/09.
PT New tapetum- and pollen-specific promoter from tobacco, useful for
preparing male sterile plants, particularly those with seedless fruits.
XX
PS Claim 6; Page 65-66; 74pp; German.
XX
XX This invention describes a novel nucleic acid (I) comprising a promoter
that is specific for tapetum and pollen. The invention also describes (a)
expression system containing at least one (I); (b) nucleic acid construct
(II) containing (I) plus at least part of an expressible nucleic acid
(III); (c) vector containing (I), the system of (a), (II), (d) cells,
particularly plant cells, containing (I); the system of (a), (II) or the
vector of (c); (e) plants containing cells of (d); (f) seeds from the

pat
parent

CC cloning a promoter that is functionally homologous with (I). Constructs
 CC containing (I) and an invertase-encoding sequence are used to produce
 CC male-sterile plants (by co-suppression or antisense techniques) for
 CC preparation of hybrids, while constructs that contain (I) and a
 CC heterologous invertase-encoding sequence are used to prepare restorer
 CC plants (which allow propagation of the male-sterile plants). Seeds from
 CC these plants are used for in vitro embryogenesis of haploid or (double)
 CC diploid plants, and the new plants particularly produce seedless fruits.
 CC (I) can also be used to prepare transgenic plants that show increased or
 CC reduced production of endogenous materials, e.g. of plant hormones or
 CC proteins involved in provision of energy to developing tissue. (I)
 CC provide high level expression in a tissue- and time-specific manner, and
 CC do not require exogenous stimuli (e.g. temperature or chemicals) for
 CC regulation
 CC
 XX SQ Sequence 4135 BP; 1212 A; 774 C; 825 G; 1318 T; 0 U; 6 Other;

Query Match 99.6%; Score 997; DB 4; Length 4135;
 Best Local Similarity 99.6%; Pred. No. 8.6e-198;
 Matches 997; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AATATAGACTTTTGAATTAATAATTAATATGATGACAAAGATTAAACCTAGTTAAT 60
 DB |||||
 QY 61 GATATTAATCTCAGATCTATTAATCAATTAACAGCAATCAGGTCATAGCAGCGTTGAGAG 120
 DB |||||
 QY 1197 GATTAATACTTCAGATCTATTAATCAATTAACAGCAATCAGGTCATAGCAGCGTTGAGAG 1356
 DB |||||
 QY 121 AAGATTAAATGTGATGTYCATTCAATATTTCAAGATCATTAATGATAGGGGAATATCAAG 180
 DB |||||
 QY 1257 AAGATTAAATGTGATGTYCATTCAATATTTCAAGATCATTAATGATAGGGGAATATCAAG 1316
 DB |||||
 QY 191 CAATTAATAAGATTAATGCGATTAATAGTAATAGGAGATGATTCACCCATATTTGAA 240
 DB |||||
 QY 1317 CAATTAATAAGATTAATGCGATTAATAGTAATAGGAGATGATTCACCCATATTTGAA 1376
 DB |||||
 QY 241 TGAGTGTGATGATCTCTCTTTTTCACCAATGATGATGATGATGATGATGATGATGATG 300
 DB |||||
 QY 1377 TGAGTGTGATGATCTCTCTTTTTCACCAATGATGATGATGATGATGATGATGATG 1436
 DB |||||
 QY 301 GACCCCTTCGAGTCTAATGAAAAAGATGGAATAGTATGATGATGATGATGATGATGATG 360
 DB |||||
 QY 1437 GACCCCTTCGAGTCTAATGAAAAAGATGGAATAGTATGATGATGATGATGATGATGATG 1496
 DB |||||
 QY 361 AGGTAGTGTGATGATCTCTCTTTTTCACCAATGATGATGATGATGATGATGATGATGATG 420
 DB |||||
 QY 1497 AGGTAGTGTGATGATCTCTCTTTTTCACCAATGATGATGATGATGATGATGATGATGATG 1556
 DB |||||
 QY 421 AGAATATTACATCCCT 480
 DB |||||
 QY 1557 AGAATATTACATCCCT 1616
 DB |||||
 QY 481 TCAATCTTAAGTACATACACCAAGATATTCATTAATTAATTAATTAATTAATTAATTAAT 540
 DB |||||
 QY 1617 TCAATCTTAAGTACATACACCAAGATATTCATTAATTAATTAATTAATTAATTAATTAAT 1676
 DB |||||
 QY 541 TATAAAATAGCTGTTAGCACTCGACCTCGGTGATGATGATGATGATGATGATGATGATG 600
 DB |||||
 QY 1677 TATAAAATAGCTGTTAGCACTCGACCTCGGTGATGATGATGATGATGATGATGATGATG 1736
 DB |||||
 QY 601 TGTCAATTACTAATCGACCTCGATTACATCTTTCTACGATGATGATGATGATGATGATGATG 660
 DB |||||
 QY 1737 TGTCAATTACTAATCGACCTCGATTACATCTTTCTACGATGATGATGATGATGATGATGATG 1796
 DB |||||
 QY 661 TTAATGAAAGCAGATTTTGACCCATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
 DB |||||
 QY 1797 TTAATGAAAGCAGATTTTGACCCATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1856
 DB |||||
 QY 721 TGGCTCTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGTAATTTATTTAAGA 780
 DB |||||
 QY 1857 TGGCTCTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGTAATTTATTTAAGA 1916
 DB |||||

QY 781 ATAGTGTGTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 840
 DB |||||
 QY 1917 ATAGTGTGTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1976
 DB |||||
 QY 841 CTTAGTAACATATATATCAAGGAATGGTGTGTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 900
 DB |||||
 QY 1977 CTTAGTAACATATATATCAAGGAATGGTGTGTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 2036
 DB |||||
 QY 901 AATATAGAGATTAGATCATCAATTAACAAAGGTTATCGTGGAGTGTGTAACAGAGCGGACC 960
 DB |||||
 QY 2037 AATATAGAGATTAGATCATCAATTAACAAAGGTTATCGTGGAGTGTGTAACAGAGCGGACC 2096
 DB |||||
 QY 961 TATGTTGTTATATAGAGGTTGTCACCCACTACTAGAAATCCGG 1001
 DB |||||
 QY 2097 TATGTTGTTATATAGAGGTTGTCACCCACTACTAGAAATCCGG 2137
 DB |||||

RESULT 4
 ADL17884
 ID ADL17884 standard; cDNA; 158001 BP.
 XX
 AC ADL17884;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human phosphotyrosyl phosphatase activator, PTPA, gene.
 XX
 KW Human; ds; antisense; phosphotyrosyl phosphatase activator; PTPA;
 KW hyperproliferative disorder; developmental disorder; infection;
 KW inflammation; tumour; gene.
 XX
 OS Homo sapiens.
 XX
 PN US2004023906-A1.
 XX
 PD 05-FEB-2004.
 XX
 PF 01-AUG-2002; 2002US-00211179.
 XX
 PR 01-AUG-2002; 2002US-00211179.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Dean NM, Dobie KW;
 XX
 DR WPI; 2004-132607/13.
 XX
 PT New antisense compound targeted to a nucleic acid molecule encoding
 PT phosphotyrosyl phosphatase activator, for modulating expression of
 PT phosphotyrosyl phosphatase activator or treating hyperproliferative
 PT disorders.
 XX
 PS Example 15; SEQ ID NO 11; 131pp; English.
 XX
 CC The invention relates to a compound 8-80 nucleobases in length targeted
 CC to a nucleic acid molecule encoding phosphotyrosyl phosphatase activator
 CC (PTPA), that specifically hybridises with the nucleic acid molecule
 CC encoding phosphotyrosyl phosphatase activator and inhibits the expression
 CC of phosphotyrosyl phosphatase activator, i.e. an antisense
 CC oligonucleotide. Also included are a composition comprising the compound
 CC and a pharmaceutical carrier or diluent, a method of inhibiting the
 CC expression of phosphotyrosyl phosphatase activator in cells or tissues, a
 CC method of treating an animal having a disease or condition associated
 CC with phosphotyrosyl phosphatase activator and a method of screening for
 CC an antisense compound. The disease or condition is a hyperproliferative
 CC disorder or developmental disorder. The compound, particularly the
 CC antisense oligonucleotide is useful in modulating the function of nucleic
 CC acid molecules encoding phosphotyrosyl phosphatase activator. The
 CC antisense compound can also be used as research tools and diagnostics. It
 CC can also be used as tools in differential and/or combinatorial analyses
 CC to elucidate expression patterns of a portion or the entire complement of
 CC genes expressed within cells and tissues. The compound can also be used
 CC for treating diseases or conditions associated with phosphotyrosyl

CC	phosphatase activator, preferably hyperproliferative disorder or
CC	developmental disorder. The compound can also be used as prophylaxis,
CC	e.g. to prevent or delay infection, inflammation or tumour formation. The
CC	present sequence is the human PRPA gene which is a target for the
CC	antisense oligonucleotides of the invention.
XX	
SQ	Sequence 158001 BP; 34764 A; 40497 C; 41696 G; 40433 T; 0 U; 611 Other;
	Query Match 5.6%; Score 55.6; DB 12; Length 158001;
	Best Local Similarity 42.6%; Pred. No. 0.085;
	Matches 346; Conservative 1; Mismatches 460; Indels 5; Gaps 1;
QY	117 AGAGAAAGATTAAATGTGATGTGCATTCAAATATTTCAGAGATCAATTAATGATAGGGGAATAT 176
DB	22124 ATATATAATATAATATATATATATATAATATAATATAATATAATATAATATAATATAA 22183
QY	177 CAAGCAATAAATAACGATAAATGGCATTAAAGTAAATAAGGAGAATGATTCACCCCAATAT 236
DB	22184 TAATATATAATAATAATAATAATATAATATAATATAATATAATATAATATAATAATAT 22243
QY	237 TGAATGAGGTGGATGATCTCTTTTGACAATGATGATGGGCAAACTACTAGATG 296
DB	22244 ATAAATAAAAAATAATAATATAAATAATATAAATAATATAATATAATATAAATAAT 22303
QY	297 TTGGGACCCCTTCFCGGATCTAATGAAAAAGATGGATAGTAGAGATAATCGAATCTCTTT 356
DB	22304 ATAAATAATAATAATAATATAATATAATAATAATAATAATAATAATAATAATATAT 22363
QY	357 AGAAAGGTAGTGAATGCTCTTTTATCTAGAGAGAAAGTCGCTTTTCAAGAAGATATTTTA 416
DB	22364 ATAAATATATAAAATATAATAATATAATATAATAAATAATAAATAATAATAATATA 22423
QY	417 TCAGAGAAATATTACATCCCCCTCTCCCTATCTCTTTTCTATTTATATGGGACATCC 476
DB	22424 TTATATAATAATATAAATAATAATAATAATATAATATAATAATAATAATAATAATAA 22483
QY	477 TCAATCAATCCTAAAAAGTACATACACCAAGAATATTTCAATAAAAAATTTTTTTGAATAT 536
DB	22484 TATATATTATATAATAATATAAATAATAATAATAATATAATATAATAATAATAATAAAT 22543
QY	537 CTATTTAABAAACTAGCTGTTAGCACTCGACCTCGGTCGGTCGYATTGACTACTCGGTTACGA 596
DB	22544 ATATAATATATATTATATAAATAATAATAAATAATAATAATAATAATAATAATAATAA 22598
QY	597 GCCTGCTCATTTACTAATCGACCTCGATTACATCACTTTCTACGATAGCTGCTTCATGCA 656
DB	22599 ATATATAAAATAATAATAATATAATATAATAATAATAATAATAATAATAATAATAA 22658
QY	657 AATCTTAATGAAGAGAGATTTTGACCCCTACAAATAATATGACAAATTTGCTTCCAAAGAA 716
DB	22659 AATATACAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 22718
QY	717 AACATGGCTTTATAGTGAATAATCGTTAGACTGTTTATAGAAGATCTCGAATTTATTTAT 776
DB	22719 AATATATAATAATAAAAAATATACAAATATAATATAATAATAATAATAATAATAAT 22778
QY	777 AAGAAATAGTGTTTTTTCTTTTCTTTTCTATCTCAAGGAGTAAAGCAACCAATGAATAGAA 836
DB	22779 ATAAAAATATAAATAATAATAATAAAAAAATAACAATATAATAATAATAAAAAATAT 22838
QY	837 AAGGCTTTAGTAAGTATATATCAAGGAATGGTGTCTTTTCTTTAAATATGGTAAAAAT 896
DB	22839 AATATATAATATAAAAAATATAATAATAATAATAATAATAATAATAATAATAATAACAT 22898
QY	897 TGTGAATATAGAAGATTAGATCAATTAACAAA 928
DB	22899 ATATAATATAAATATAATAATAATAATA 22930
RESULT 5	
ABZ10246/c	
ID	ABZ10246 standard; DNA; 8056 BP.
XX	

AC	ABZ10245;
XX	
XX	16-JAN-2003 (first entry)
DE	
XX	Haematopoietic cell proliferation disorder related DNA sequence #386.
XX	
KW	Human; haematopoietic cell proliferation disorder; cytostatic;
KW	gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW	cytosine methylation state; gene; ds.
XX	
XX	Homo sapiens.
OS	
XX	WO200277272-A2.
PN	
XX	
PD	03-OCT-2002.
XX	
PF	26-MAR-2002; 2002WO-EP003401.
XX	
PR	26-MAR-2001; 2001US-0278333P.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
XX	Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI	Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI	Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI	Schwabe I, Ziebarth H;
XX	
DR	WPI; 2003-018942/01.
XX	
PT	Detecting and differentiating between hematopoietic cell proliferative
PT	disorders, comprises contacting a target nucleic acid with a reagent that
PT	distinguishes between methylated and non-methylated CpG dinucleotides.
XX	
XX	Claim 28; SEQ ID NO 386; 117pp; English.
PS	
XX	
CC	The present invention describes a method for detecting and
CC	differentiating between haematopoietic cell proliferative disorders
CC	associated with at least 1 gene and/or their regulatory regions in a
CC	subject. The method comprises contacting a target nucleic acid in a
CC	biological sample obtained from the subject with at least 1 reagent,
CC	which distinguishes between methylated and non-methylated CpG
CC	dinucleotides within the target nucleic acid. AB209861 to AB211118
CC	represent specifically claimed nucleotide sequences from the present
CC	invention. Oligonucleotides from the present invention can be used: for
CC	differentiating between healthy haematopoietic cells and proliferative
CC	disorder haematopoietic cells; for differentiating between acute
CC	lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC	determining the cytosine methylation state and/or single nucleotide
CC	polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC	related sequences and their complements; and as primers for the
CC	amplification of haematopoietic cell proliferation disorder related DNA
CC	sequences. The nucleotide sequences from the present invention can also
CC	be used for detecting a predisposition to, differentiation between
CC	subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC	haematopoietic cell proliferative disorders. The present method enables a
CC	highly specific classification of haematopoietic cell proliferative
CC	disorders allowing for improved and informed treatment of patients
XX	
SQ	Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

	Query Match	5.5%	Score 55.2	DB 8	Length 8056
	Best Local Similarity	42.8%	Pred. No. 0.067		
	Matches 395	Conservative 1	Mismatches 514	Indels 12	Gaps 2
Qy	3	TATAGACTTTTGATTTAAATTAATTAATTTGTTATGACAAAGGATTAACCTAGTTAATGA	62		
Db	2610	TATAAATTTTTTTTTTCATTTATTTTTATTTTTTAAAAAATTTTTATTAAAAATTTTAAAT	2551		
Qy	63	TAATAACTTCAGATCATATATCATTAATTAACAGCAATCACGGTCATAGCAGGTTGAGAGAA	122		
Db	2550	TTTTAAATTTAATTTATAAATTTAATTTATTTTTTTTTTTTTCATTAATATTTTAAATAT	2491		
Ov	123	GATTAAATGATGATGYCATTCATATTTTCAAGATCATTAATGATGAGGGGAATATCAAGCA	182		

RESULT 5
ABZ10246/C
ID ABZ10246 standard; DNA; 8056 BP.
XX

Db	2490	TTTAAACAAAATAATAAAATAATATAAAATAAAAAATAAAAAATAATTTTAAAAA	2431
Qy	183	ATAAATAACGATAAATGGCAATTAAAGTAAATAAGGAGAATGATTCACCCCAATATTGAATG	242
Db	2430	TAATAAATATATTTTTTTTTTAAATTTTCAAAAAATAATAAATTAATAAATATTATTATA	2371
Qy	243	AGTGGATGATCTCTTTTTTGACAATGATGATGGGCAATATCTAGAAATGTCGGGA	302
Db	2370	AAAAATAAAAAATATATTTTAAAAATTAATAAAAAATTTATTAAATACAAAAATTAATAAAA	2311
Qy	303	CCCTTCTCGGATCTAATGAAAAAGTATGGAATAGTAGATAATCGAATCTCTTAGAAG	362
Db	2310	TTATTTT---AAAAATAATAATTAATAAATTAATAATTTTAAATTTATTATTAAATA	2255
Qy	363	GTAGTGAATGTCTTTTATCTAGAGAGAAAGTCGTCTTTTCAAGAATATTTTTATACAG	422
Db	2254	TTAAAAATAAAATTTTATTAATAAATTAATTTTTTTTTTTTTTTTATTTTTTTTCAATT	2195
Qy	423	AATATTACATCCCCCTCTCCCTCATCTCTTTTCTATTTATATGAGACATCTCCTCAATC	482
Db	2194	TTTTAAATTTTAAATAAAATTAATTTTTTTTTTTTTTTTTTTTATTTTATTTTT	2143
Qy	483	AATCCTAAAAAGTACATACACAAGAATATTCAATAAAAAATTTTTTGAATTTCTAATA	542
Db	2142	AATAAAATAAAAAAATTAATAAATTAATAAATAATTAATAATTTATTATTATAATATT	2083
Qy	543	TAAAAACTAGCTGTTAGCACTCGACCTCGGTGGTATTTGACTACTCGGTACGAGCCCTG	602
Db	2082	TAAATATTTTTTAATAAAATTTTTTTAAAAATAAAAAATATTTAAAAATATTTTTATTTT	2023
Qy	603	TCATTTACTAATCGACCTCGATTACATCACTTTCTAGCATACTGCTTCATGTCAAATCTT	662
Db	2022	AAATATATTTAAATTTTTTATTAATTTTAAATAAAAAATTAATAATATTTTATTTTAA	1963
Qy	663	AATGAAAGCAGATTTTGGACCCATCAATAATATGACAAAATGCTTCCAAAGAAAACATG	722
Db	1962	AAAAATATATTTTTTTTTTATTTATATATAATTAATTTATTTTTTTTTTATAAAATTTT	1903
Qy	723	GCTCTTATAGTAAATATCGTTAGACTGTTATAGAAAAGATCTGAATTTATTTATAAGAA	782
Db	1902	ATTTTATATATAAAAAATAAAAAATAATTAATTAATTTATTTTTTTTTTTTTTTTTT	1843
Qy	783	AGTGTTTTTTCTTTTCTTTTCATATCTAGGAGTAAAGCAACCATGGAATAGAAAAGCT	842
Db	1842	TTTTTTTTTAAACAAAATTTTATTTTATTTAATTAATAAATAATAAATTAATAATTTT	1783
Qy	843	TAGTAACTATATATCAAGGAATCGTGTTTTTTCTTTAAATATGATAAAAAATTTGTGA	902
Db	1782	TTAATTTTTTAAATTTTAAAAAATTTTTTTTATTTTAAATTTTTTTTAAATAATTAATA	1723
Qy	903	TATAGAAATTAGATCAATTAA	924
Db	1722	TATATATTATTAATTAATAATTA	1701

RESULT 6	
ABQ66998/c	
ID	ABQ66998 standard; DNA; 37515 BP.
XX	
XX	ABQ66998;
XX	
XX	
DT	28-AUG-2002 (first entry)
XX	
DE	Human angiogenesis associated polynucleotide SEQ ID NO 28.
XX	
XX	Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW	inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcerds;
KW	macular degeneration; inflammatory bowel disease; Crohn's disease;
KW	antirheumatic; antiarthritic; antiabietic; antipsoriatic;
KW	antiarteriosclerotic; ds.
XX	
XX	
OS	Homo sapiens.

XX WO200246454-A2.
PN
XX
XX PD
XX 13-JUN-2002.
XX
XX PF 06-DEC-2001; 2001WO-EP014320.
XX
XX PR 06-DEC-2000; 2000DE-01061338.
XX
XX PA (EPIG-) EPIGENOMICS AG.
XX
XX PI Schacht O;
XX
XX DR WPI; 2002-500450/53.
XX
XX PT New nucleic acid fragments from chemically treated angiogenesis-associated genes, useful for determining methylation status, e.g. in diagnosis or treatment of cancer.
XX
XX
XX Claim 1; SEQ ID NO 28; 41pp + Sequence Listing; German.
XX
XX The invention relates to a nucleic acid (I) comprising a segment of 18 bases of chemically pretreated DNA of angiogenesis-associated genes (II) having sequences (ABQ66971-ABQ67178) or their complements. (I), also related oligomers, are used to evaluate the methylation status and/or single-nucleotide polymorphisms, in angiogenesis-related genes, for diagnosis and treatment of eye diseases, proliferative retinopathy, CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis, CC diabetic retinopathy, macular degeneration caused by neovascularisation, CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and CC Crohn's disease. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format CC directly from WIPO at ffp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 37515 BP; 11428 A; 425 C; 7165 G; 18497 T; 0 U; 0 Other;
SQ

Query Match 5.5%; Score 54.8; DB 6; Length 37515;
Best Local Similarity 46.4%; Pred. No. 0.1;
Matches 245; Conservative 1; Mismatches 278; Indels 4; Gaps 2;

QY 389 AAAGTCGTCTTTTCAAAGAATATTTTTATCAGAGAAATTATCATCCCCCTCTCCCTAT 448
DB 12156 AAATACTTACATAAACCACACATATATCATATTAATAACTTAATTTTAAAT 12097
QY 449 CTCCTTTCTATTTATATGGGACATTCCTCAATCAATCCTAAAAGTACATACACAAGAA 508
DB 12096 TTCTCTAATTTAATTATCCTATATATAAATTTAATTTTCCAATACAAATCAAAAA 12037
QY 509 TAATCAATAAATAATTTTTGAATTTCTATTATAAAACATAGCTGTTAGCACTCGACC 568
DB 12036 TAATAAATATAATTTTAAATTTTAAATTAATAAATAAATAAATAAATAAATAAATAA 11977
QY 569 TCGGTGCVATTGACTACTCGGTTACGAGCCCTGTCAITTTACTTAATGACCTCGATTACA 628
DB 11976 TAAATTCATATTAATTTTAATTTAACGATAT-TTATTTTACTTAACATAAATCTC 11918
QY 629 TCACITTTACGATACCTGTCATGTCAAATCTTAATGAAGACGATTTTGACCCATACA 688
DB 11917 TTACAATATAAAATATTACTTTA---CTACATAATCAATTTACACCTCAACGTAAC 11861
QY 689 ATAATATGACAAATTTGCTTCCAAAGAAAAATGGCTCTTATAGTGAATATCGTTAGAC 748
DB 11860 TTAAATTTAAANAATACATATATAAATAACATTTCTACTATCTTAAACATTTTTTTA 11801
QY 749 TGTTATAGAAGATCGTAATTTATTTATAGAATAGTGTTTTTTCTTTCTTTTCATAT 808
DB 11800 ACCTCCTTAACCATACATTTTAAATTTTAAACATAAATAAACAATCTTTTTTTTAAAT 11741
QY 809 CTAAAGGATTAAGACCAACCATGAATAGAAAAAGCTTAGTAACTATATATCAAGGAATGG 868
DB 11740 TAAAAAAAAAAAACCTTACTACCCAAAAATTAANAATAAANAATAAANAATAAANAATAA 11681
QY 869 GTTTTCTTTTAAATATGATATAAATTTGTGAATATAGAAGATTAGA 916

Db 11680 ACTTAATATATTAATTAATAAAGATTATTTTAAAAAACAACAACTAAAA 11633
ABL34175/c
RESULT 7
ID ABL34175 standard; DNA; 113515 BP.
XX
AC ABL34175;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 2148.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriasis;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
FN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007537.
XX
PF 30-JUN-2000; 2000DE-01032529.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.
XX
XX Claim 1; SEQ ID NO 2148; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention
XX
XX Sequence 113515 BP; 35172 A; 1174 C; 22520 G; 54649 T; 0 U; 0 Other;
XX
XX Query Match 5.4%; Score 54.4; DB 6; Length 113515;
XX Best Local Similarity 46.1%; Pred. No. 0.14;
XX Matches 178; Conservative 1; Mismatches 207; Indels 0; Gaps 0;
XX
XX 453 TTTTCTATTTATATGGACATCTCTCAATCAATCTTAAAGTACATACCAAGATATT 512
XX 4047 TATTATAAAAAATAAACTTTTACAAACTTATCTTACAAACATTACTCAAATAATTAT 3988
XX
XX 513 CAATAAATATTTTTTGAATATTTCTATTATAAACTAGCTGTAGCACTCGACCTCGG 572
XX
XX 3987 ATATAAATTACTTATCAAAAAAATTTCATAAAAAATTATTAATAAAAAATTAACTA 3928
XX
XX 573 TCGYATTTGACTACTCGGTGTACGAGCCCTGCTATTTACTAATCGACCTCGATTACATCAC 632
XX
XX 3927 TCATTAATAACATTTTAAAAAATATCTTTAAATACTACTAATTACTAAATCTA 3868

QY 633 TTTCTAGTACTGCTTCTCATGTCAATCTTAATGAAGCAGATTTTGACCCATACAATAA 692
Db 3867 AAACAAAAATACAAATTAATAAATTTTAAATAAATACATATACATATACTATAACTTTA 3808
QY 693 TATGACAAAATTGCTTCCAAAGAAAACATGGCTCTTATAGTGAAATATCGTTAGACTGTT 752
Db 3807 ACAAAACITTTCTAAATTAATAACAAAACAATATATATATATTTAATTAACATTCC 3748
QY 753 ATAGAAAGATCTGAATTTATTTATTAAGAATAGTGTGTTTTTCTTTTCTTTTCTATCTAA 812
Db 3747 AAAAAATATACCTATTTTATATAAATAATATTTATTTATATAAATAATTTTAAACAATCCAC 3688
QY 813 GGAGTAAAGCAACCATGATGAAA 838
Db 3687 AAATCTCAAAACCATTAATTATATA 3662
RESULT 8
ID ABL34511 standard; DNA; 8201 BP.
XX
AC ABL34511;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human metastasis associated gene SEQ ID NO: 64.
XX
KW Metastasis associated gene; cytostatic; gene therapy; cancer;
KW cytosine methylation; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200177376-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-EP003970.
XX
XX 06-APR-2000; 2000DE-01019058.
XX
XX 07-APR-2000; 2000DE-01019173.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-010922/01.
XX
XX New nucleic acid derived from chemically treated metastasis genes, useful
XX for diagnosis of cancers by analysis of cytosine methylation, also for
XX treatment.
XX
XX Claim 1; SEQ ID NO 64; 23pp + Sequence Listing; English.
XX
XX The present invention provides a number of human metastasis associated
XX genes which are modified by cytosine methylation. The sequences can be
XX used in the diagnosis and treatment of cancer. The present sequence is
XX one of the genes of the invention
XX
XX Query Match 5.3%; Score 52.6; DB 6; Length 8201;
XX Best Local Similarity 53.1%; Pred. No. 0.23;
XX Matches 112; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
XX
XX 689 ATAAATATGACAAAATGCTTCCAAAGAAAACATGGCTCTTATAGTGAATATCGTTAGAC 748
XX 4834 ATAAAAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4893
XX 749 TGTATAGAAAGATCTGAATTTATTTATAAGAATAGTGTGTTTTTCTTTTCTTTTCTATAT 808
XX 4894 TGTATAGATATGAGGAATGTTTTTATGTTGATATTTTTTTCGTAATTTTATATATGT 4953

QY 809 CTAAGGAGTAAGCAACCATGATAGAAAAGGCTTAGCTACTATATATCAAGAAATGCT 868
 DB 4954 TGGGATTGTAAGTTATATAGTATATATTTGTTGAAGTTATTTTATTTTAAAGGAGTAAT 5013
 QY 869 GTTTTCTTTTAAATATGGATAAAATTTGT 899
 DB 5014 TTAGTTTTTTTATTTTGAATTAAGATTTGCT 5044

RESULT 9
 ABL70540
 ID ABL70540 standard; DNA; 8201 BP.
 AC ABL70540;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Chemically treated cell signalling DNA sequence complementary to#215.
 XX
 KW Cell signalling; cytosine methylation; cell signalling disease; cancer;
 XX tumour; cytostatic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200202807-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-EP007471.
 XX
 PR 30-JUN-2000; 2000DE-01032529.
 XX
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-154758/20.
 XX
 PT Nucleic acid, useful for diagnosis and therapy of diseases associated
 PT with cell signaling e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with cell signaling.
 XX
 PS Claim 1; SEQ ID NO 430; 24pp + Sequence Listing; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of at least
 CC 18 bases of a segment of chemically pretreated DNA of genes associated
 CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or PNA-oligomers for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling. Note: The
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office
 XX
 SQ Sequence 8201 BP; 1609 A; 455 C; 2470 G; 3667 T; 0 U; 0 Other;
 Query Match 5.3%; Score 52.6; DB 6; Length 8201;
 Best Local Similarity 53.1%; Pred. No. 0.23;
 Matches 112; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 689 ATAATATGACAAATGCTTCCAAGAAACATGGCTTCTTATAGTGAATATCGTTAGAC 748
 DB 4834 ATAAAAATAAAAATATAATAAAAAATACGTTAATAGTTAGTGAATATTTATT 4893

QY 749 TGTATAGAAAGATCTGAATTTATTTATAGATAGTGTCTTTCTTTCTTTCTATAT 808
 DB 4894 TGTATAGATATGAGGAATGTTTATCTGATATTTTTCGTAATTTATATGT 4953
 QY 809 CTAAGGAGTAAGCAACCATGATAGAAAAGGCTTAGCTACTATATATCAAGAAATGCT 868
 DB 4954 TGGGATTGTAAGTTATATAGTATATTTGTTGAAGTTATTTTATTTTAAAGGAGTAAT 5013
 QY 869 GTTTTCTTTTAAATATGGATAAAATTTGT 899
 DB 5014 TTAGTTTTTTTATTTTGAATTAAGATTTGCT 5044

RESULT 10
 ABZ10100/C
 ID ABZ10100 standard; DNA; 8056 BP.
 XX
 AC ABZ10100;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #240.
 XX
 KW Human; haematopoietic cell proliferation disorder; cytostatic;
 KW Gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277272-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-EP003401.
 XX
 PR 26-MAR-2001; 2001US-0278333P.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
 PI Schwope I, Ziebarth H;
 XX WPI; 2003-018942/01.
 XX
 PT Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 XX
 PS Claim 28; SEQ ID NO 240; 117pp; English.
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ1118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia, as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative

CC disorders allowing for improved and informed treatment of patients

XX
SQ Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;

Query Match 5.2%; Score 52; DB 8; Length 8056;
Best Local Similarity 42.6%; Pred. No. 0.31;
Matches 393; Conservative 1; Mismatches 516; Indels 12; Gaps 2;

QY 3 TATAGACTTTTGATTAAATAAATTAATGATGTATGCACCAAGGATTAACCCTAGTTAATGA 62
 |||||
Db 2610 TATAAATTTTTTCGGTTATTTTATTTTATTTTAAAAATTTTATTTAAAATTTTAAAT 2551
 |||||
QY 63 TAATAACTCAGCATCATATCAATTAACAAGAATCACCGTCATAGCAGCGTTGAGCAA 122
 |||||
Db 2550 TTITTAATTTATTTAATTTAATTTATTTTATTTTATTTTTCGTATATTTAATTTAT 2491
 |||||
QY 123 GAATAATGTGATGYCATCAATATTTCAAGATCANHTAATGATAGGGGATATCAAGCA 182
 |||||
Db 2490 TTTAAACGAAATTAATAAAAAAAAAATATAATAATAAATAAATAAATAAATAAATAAATA 2431
 |||||
QY 183 ATAAATAACATAAATGCCATTAAGCTAAATAGGAGATGATTCACCCAATATTGAATG 242
 |||||
Db 2430 TAAATAAATTTTATTTTATTTTAAATTTTCGAAATAATAAATAATTAATAATTTATA 2371
 |||||
QY 243 AGTGGAGATCTCTCTTTTGCAATGATGAATGAGGGCAAACTACTAGAATGTGGGA 302
 |||||
Db 2370 AAAATAAAAAATATATTTAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 2311
 |||||
QY 303 CCCTCTCGGATCTAATGAAAAAGATGGAATAGTAGATATCGAATCTCTTTAGAAC 362
 |||||
Db 2310 TTAATTT---AAAAATAATAAATTAATAAATAAATAAATAAATAAATAAATAAATA 2255
 |||||
QY 363 GTAGTATGTCCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAATAATTTTATFCAGAG 422
 |||||
Db 2254 TTAATAAATAAATAAATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 2195
 |||||
QY 423 AAATATACATCCCCCTCTCCCCTATCTCTTTTCTATTTATATGAGCATCTCCTCAATC 482
 |||||
Db 2194 TTTTAAATTTTAAATAAATAATATATTTTATTTT-----ATTTTATTTT 2143
 |||||
QY 483 AATCCTAAAAAGTACATACACCAAGATATTTCAATAAAATATTTTTCGAATATCTAPTA 542
 |||||
Db 2142 AATAAATAAATAAATAATTTAATAAATAATTTAATAAATAATTTATAATTTATTTAATAAT 2083
 |||||
QY 543 TAAAACTAGCTGTAGACTCGACCTCGTGGYTATGACTACTCGGTAGCAGCCCTG 602
 |||||
Db 2082 TAAATATTTTAAATAAATAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATA 2023
 |||||
QY 603 TCATTTACTAATCGACCTCGATPACATCACTTTTACGATATCGTCTCATGTCAAACTCT 662
 |||||
Db 2022 AAATATATTTAAAAATTTTATTAATTTTAAATAAATAAATAAATAAATAAATAAATAA 1963
 |||||
QY 663 AATGAAGCAGATTTTGACCCATACAAATATATGACAAATTTGCTTCBAAGAAAAATG 722
 |||||
Db 1962 AAAATATATATTTTATTTATTTATTTAATAATTTAATTTTATTTTATTTTAAATAATTT 1903
 |||||
QY 723 GCCTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGAATTTATTTATAAGAAT 782
 |||||
Db 1902 ATTITTATATAAABAATAAATAAATAAATAATTAATAAATAATTTTATTTTATTTT 1843
 |||||
QY 783 AGTGTTTTTTCTTTCTTTTCATCTAAGGAGTAAGCAACCATGAATAGAAAAAGCT 842
 |||||
Db 1842 TTTTTTAAACGAAATTTTATTTTATTTAATAAATAAATAAATAAATAAATAAATAAATA 1783
 |||||
QY 843 TAGTAATATATATCAAGGAATGGTGGTGTCTTTTCTTTAATATGATTAATAAATTTTGCA 902
 |||||
Db 1782 TTAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1723
 |||||
QY 903 TATGAAGATTAGATCAATTA 924
 |||||
Db 1722 TATATATTTATTTAATAAATA 1701

RESULT 11

ABL326L5
ID ABL326L5 standard; DNA; 13594 BP.

XX AC ABL326L5;
XX DT 26-MAR-2002 (first entry)
XX DE Human immune system associated gene SEQ ID NO: 588.
XX KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiaertherosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; antiviral; ophthalmological;
KW rheumatic; antiarthritis; antididiabetic; antipsoriasis;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
ds.
OS Homo sapiens.
XX WO200200928-A2.
XX PN 03-JAN-2002.
XX PD 02-JUL-2001; 2001WO-BPO07537.
XX PF 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PA (EPIC-) EPGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-130909/17.
XX PT Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation.
XX PS Claim 1; SEQ ID NO 588; 32pp + Sequence Listing; German.
XX CC The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/alcerative bowel diseases. The present sequence is a gene of the invention
XX SQ Sequence 13584 BP; 3905 A; 112 C; 2520 G; 7047 T; 0 U; 0 Other;

Query Match 5.2%; Score 52; DB 6; Length 13584;
Best Local Similarity 46.6%; Pred. No. 0.34;
Matches 264; Conservative 1; Mismatches 296; Indels 5; Gaps 3;

QY 1 AATATAGACTTTTGATTAATAATTAATTTATTTATGATGATCAAGATTAACCTAGTTAAT 60
 |||||
Db 12840 ATTAGTGAGAAATGTAATTTTGTGTGATATTTTGAAGAGAAAAATGSGGAATATT 1289
 |||||
QY 61 GATAATAACTTCAGATCTATTAATCAATTAACAGCAATCACCGTCAATAGCAGGTGAGAG 120
 |||||
Db 12900 TTGTAGATAAAGATTTATTTATTAAGTTATAGTAATTTATGTTATATGTTTGTGT 1295
 |||||
QY 121 AAGATTAAATGTGATGTYCATTCAT--ATTTCAAGATCATTAATGATAGGGAATATC 177
 |||||
Db 12960 GTGGGAATAATAATATTAATGGAATAGATAGAAAGTTTGAATAATGATTGAAATTT 1301
 |||||
QY 178 AAGCAATAAATACGATAAATGCGATTAAAGTAAATAAGGAGAAATGATTCACCAATATT 237
 |||||
Db 13020 GATATATAATAGAGGTTACGTGTTATTTAGAAAAAGATGATTTTAAATAAATGGTGT 1307
 |||||

Qy	238	GAATGAGGTGGATGATCTCTCTTTTGCAANTGATGATGGGCAAACTACTAGAACTG	297
Db	13080	GAGATAATTGGTTATTTATGTAGAAATAAAATTAGATTTTATTTTATAGTAAATTA	13139
Qy	298	TGGGACCCCTCTCCGATCTAAATCAAAAAAGTAGGAATAGTAG-ATAATCGAATCTCTTT	356
Db	13140	TATATATAATTTTTTAAATGATTAAGGTTTAAATAGAAAAAGTAAATTTTAAATTTTT	13199
Qy	357	AGAAAGGTAGTGATTCCTCTTTTATCTAGAGAGAAAGTCGTCTTTTCAAGATATTTTTTA	416
Db	13200	AGAAGAAAAAGTATGTTGTATAATTTTGGATAAAAGTATTTTTTAAATAGTTATATA	13259
Qy	417	TCAGAGAAATATACATCCCCCTCTCTCCCTATCTCTTTTCTTATTATA-TGGGACATTC	475
Db	13260	AATGCTGATATATAAAGGAAACGATTGATAAATTTATTTTATTAAAAATAAAAATG	13319
Qy	476	CTCAATCAATCTTAAAGTACATACACCAAGATATTCGAATAAATATTTTTTGAATAT	535
Db	13320	AGATATTAATAAAGAGTGAAGATGAAGTTTAGATTGAATGAAGTATTTGTTAGTTAA	13379
Qy	536	TCATATTATAAAACTAGCTGTTAGCA	561
Db	13380	GTTTTAGAAAAGATTAAATATTAGAA	13405

RESULT 12
ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.
XX
XX
AC ABZ10246;
XX
XX
16-JAN-2003 (first entry)
DT
XX
XX
Haematopoietic cell proliferation disorder related DNA sequence #386.

XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Pispembrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwobe I, Ziebarth H;
XX WPI; 2003-018942/01.
DR
XX
XX Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
XX Claim 28: SEQ ID NO 386: 117bp: English.
PS

XX The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. AS3039861 to AS321118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used; for

differentiating between healthy haematopoietic cells and proliferative disorder haematopoietic cells; for differentiating between acute lymphocytic leukaemia and acute myelogenous leukaemia; as probes for determining the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) of haematopoietic cell proliferation disorder related sequences and their complements; and as primers for the application of haematopoietic cell proliferation disorder related DNA sequences. The nucleotide sequences from the present invention can also be used for detecting a predisposition to, differentiation between subclasses, diagnosis, prognosis, treatment and/or monitoring of haematopoietic cell proliferative disorders. The present method enables a highly specific classification of haematopoietic cell proliferative disorders allowing for improved and informed treatment of patients

Query Match	5.1%	Score 50.8	DB 8	Length 8056
Best Local Similarity	44.1%	Pred. No. 0.55		
Matches 410	Conservative	2	Mismatches 494	Indels 24
Gaps				
QY	11	TTTGATTAATTAATTAATTCGTATGACAAAGGATTAACCTAGTTAATGATAAATAACT	70	
DB	1359	TTTTTTATTAATAATTTATTTTTTTTATAAAATAATATGATTTTTTTTTTATTTTT	1418	
QY	71	TCAGATCTATATCAATTAACAGCAATCAGGTATAGCAGCGTTCAGAGAGATTAAT	130	
DB	1419	TTATTTTTTTTTTTAAAAAATAAAATATTTTTTTTAATAAAATATATTTAAAAA	1478	
QY	131	GTGATGYCATTCCAATATTTCAAGATCATTAAATGATAGGGGAATATCAAGCAATAATAA	190	
DB	1479	TAAATTAATTTAAATTAATTTATATAATAATAAAATATAAATATGTTAAAAAATTT	1538	
QY	191	CGATAATGGCATTAAGTAATAAGAGAGATGATTCACCCAATATGATGAGGTGGAT	250	
DB	1539	TAAATAAAATATTTTTTTAAAAAATATATTTTAAATTTTAAATTTTAAATTTAT	1598	
QY	251	GATTCCTCTTTTTGCAATGATGAATGA-----TGGGCAATACTAGAAATGTCGGAC	303	
DB	1599	AATTTTAATTTTTAAAAAATCGAAATAAATGAATGTAATAATTAATAAAATATTTTA	1658	
QY	304	CCCTCTCGGATCTAATGAAAGATGTAAGATAGTAGATAATCGATCTCTTTAGAAAGG	363	
DB	1659	ATTAAATGATGATAAAATTTTATTTTTTATTAATAATTAATAAAATTAATTAATTA	1718	
QY	364	TAGTGATGTGTTTTATCTAGAGAGAAAGTCGTCTTTTCAAAGAATATTTTTATCAGAGA	423	
DB	1719	TATATTTATAATTTTAAATTTTAAAAA--ATTAAAAATAAAAAATTTTTTAAATTA	1776	
QY	424	ATATTACATCCCGCTCTCCCTATCTCTTTTCTATTTATATGAGACATCTCCTCAATCA	483	
DB	1777	AAATAAAATTAATATGTTATTTATTTTAAATTTAAATAATAAAATTAATTTGTTTA	1836	
QY	484	ATCCTAAAAAGTACATACACCAAGAATATTCATATAAAATATTTTTTGAATATCTATTAT	543	
DB	1837	AAAAAATAAAAAAATAAAAAAATAATTTAAATTAATTTATTTTTTTTTTTTATAT	1896	
QY	544	AAAAACTAGCTGTTAGCACTCGACCTCGGTCGTATTGACTACTCGGTCACGAGCCCTGT	603	
DB	1897	AAAAATAAAATTTTTATAAAAAATAAAATAATTTAAATATATAATAATAAAAAATAT	1956	
QY	604	CATTTACTAATCGACCTCGATTAATCATCATCTTTCTCAGATCTCGTCACTCAAAATCTTA	663	
DB	1957	AATTTTTTAAAAATAAATAATTAATAATAATTTTTTA-----TTTAAAAATTTAAATA	2008	
QY	664	ATGAAAGCAGATTTTGACCCATACAAATAATACAAAAATTTGCTTCCAAAGAAAAATGG	723	
DB	2009	TTTTTAATAATTTTAAAAATAAATAATATTTTAAATATTTTTTATTTTTTAAAAATTTA	2068	
QY	724	CTCTTATAGTGAATATCGTTAGACTGTTTATAGAAAGATCTGAATTTATTTAAGATA	783	
DB	2069	TT-----AAAAATATTTTTAAATATTTATAAATAAATAATTAATAATTTTAAATAAT	2121	
QY	784	GTGTTTTTTCTTTCTTTTCAATCTAAGGAGTAAGCAACCATGAATAGAAAAGGCTT	843	

PT genes and oncogenes, useful in designing primers and probes for analyzing
 PT diseases associated with cytosine methylation state e.g. cancer.
 XX
 PS Claim 1; SEQ ID NO 417; 27pp; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and oncogenes
 CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
 CC 500 are missing from the sequence listing) sequences (Ss) and sequences
 CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
 CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
 CC probes for detecting the cytosine methylation state and/or single
 CC nucleotide polymorphisms and also to be used in an array for analysing
 CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
 CC probes can also be used in a method for ascertaining genetic and/or
 CC epigenetic parameters for the diagnosis and/or therapy of existing
 CC diseases or the predisposition to specific diseases, by analysing
 CC cytosine methylations. The parameters may be compared to another set of
 CC genetic and/or epigenetic parameters, the differences serving as basis
 CC for diagnosis and/or prognosis events which are disadvantageous to
 CC patients. The present sequence is one of the 533 genomic sequences
 CC derived from tumour suppressor genes and oncogenes. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5413 BP; 1391 A; 65 C; 1102 G; 2855 T; 0 U; 0 Other;
 Query Match 5.0%; Score 50; DB 4; Length 5413;
 Best Local Similarity 51.5%; Pred. No. 0.77;
 Matches 139; Conservative 0; Mismatches 130; Indels 1; Gaps 1;
 QY 652 TGTCAATCTTAATGAAGCAGATTTCCACCATACATATGACAAATTCCTTCCA 711
 Db 1263 TTTTATATAGAGTAAATTTATTTTCGTAGTTATTTTGTATATATATTTTAT 1322
 QY 712 AAGAAACATGGCTTATAGTGAATATCGTAGCTGTATAGAAGATCGAATTTA 771
 Db 1393 ATATATATATTTATATATATTTTATATATATATATA-TTTTATATATATTTTATA 1381
 QY 772 TTTATAGAATAGTGTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 831
 Db 1382 TATATATATAGAGTATTTTATTTTATTTTATTTTATAGAGTATTTATATATATTTT 1441
 QY 832 TAGAAAGCGCTAGTAACTATATATCAAGAAAGGTTGTTTCTTTTAAATATGGATAA 891
 Db 1442 TATATAATATATATATATTTTATTTTATAGTATTTTATATATTTATATATAGATT 1501
 QY 892 AAATTTGGAATATAGAAGATTAGATCAAT 921
 Db 1502 ATTTTATATATAGGATTTATATATTTAT 1531
 RESULT 15
 ID ABK28332/c
 XX ABK28332 standard; DNA; 11745 BP.
 AC ABK28332;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA transcription associated complementary genomic DNA #103.
 XX
 KW PNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
 KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
 KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
 KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
 KW immunological disorder; Werner syndrome; developmental disorder;
 KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
 KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
 KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
 KW angiodysplasia; congenital heart disease; HDR syndrome; gene therapy;

KW polyglutamine disorder; solid tumour.
 XX Unidentified.
 OS
 XX WO200192565-A2.
 PN
 XX 06-DEC-2001.
 XX
 PD 06-APR-2001; 2001WO-BP003973.
 XX
 PF 06-APR-2000; 2000DE-01019058.
 XX
 PR 07-APR-2000; 2000DE-01019173.
 XX
 PR 30-JUN-2000; 2000DE-01032529.
 XX
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-090046/12.
 XX
 DR New nucleic acids or oligomers, useful for diagnosing or treating
 XX diseases associated with DNA transcription, e.g. immunological disorders,
 PT Werner syndrome, psoriasis, myocardial infarction, solid tumors or
 PT cancer.
 PT
 XX Claim 1; SEQ ID NO 206; 32pp; English.
 PS
 XX The invention relates to a nucleic acid, which comprises a segment of the
 CC chemically pretreated DNA of genes associated with DNA transcription from
 CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
 CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
 CC to the chemically pretreated DNA of genes associated with DNA
 CC transcription. The set of oligomer probes are useful for detecting the
 CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
 CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
 CC diagnosing or treating diseases associated with DNA transcription
 CC (particularly with the methylation status), e.g. adenosine deaminase
 CC deficiency, viral infection, retroviral infection, Sezary syndrome,
 CC haematological disorders, immunological disorders, Werner syndrome,
 CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
 CC neurological disorders, neurodegenerative disorders, Waardenburg
 CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
 CC infarction, hypertension, angiodysplasia, erythropoiesis, congenital heart
 CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
 CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
 CC associated genomic DNA molecules of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification but
 CC was obtained in electronic format directly from the European Patent
 CC Office
 CC
 XX Sequence 11745 BP; 3183 A; 270 C; 2383 G; 5908 T; 0 U; 1 Other;
 SQ
 Query Match 5.0%; Score 50; DB 6; Length 11745;
 Best Local Similarity 47.9%; Pred. No. 0.86;
 Matches 202; Conservative 1; Mismatches 216; Indels 3; Gaps 2;
 QY 13 TGATTAAATTAATTAATTTGTATGACAAAGATTAAACCTAGTTATGATAATAACTTC 72
 Db 6199 TTAACAATATCTTTAAATAATCAATAATAATAAAAAATCGAAAAACCAATAAAAAAAA 6140
 QY 73 AGATCTATATCAATTAACAGCAATCACGGTCATAGCAGC-GTTGAGAGAAGATTAAATG 131
 Db 6139 ATATAACTATTAAATTTTAAAAAATCAATATTAATAATATCTCTAAAAATATAATAATT 6080
 QY 132 TGATGTCATTCATTAATTTTCAAGATCATTAATGATAGGGGAATATCAAGCAATAATAAAC 191
 Db 6079 TATATATCATCTAAAAAATAAAAAATAATTAATAATATATACAAATCAATAATAATAAT 6020
 QY 192 GATAAATGGCATTAAAGTAATAAGGAAATGATTCACCCATATTCAGATGAGTGGATG 251
 Db 6019 ATCTAACCAATAATACTACTTTTAAATAATAATAATAATAATAATAATAATAATAATA 5960

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 07:44:14 ; Search time 3107.21 Seconds
(without alignments)
11739.223 Million cell updates/sec

Title: US-10-009-966C-2_COPY_2155_3155
Perfect score: 1001
Sequence: 1 aaatagacttttgattaaa.....accactactagaatccgg 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	66.2	6.6	1101	9	CNS0039G
c 2	66	6.6	1101	9	CNS0039G
c 3	63.2	6.3	1101	9	CNS0039G
c 4	62	6.2	1101	9	CNS0039G
c 5	60.2	6.0	1092	9	CNS020K7
c 6	59.4	5.9	1101	9	CNS0039G
c 7	59	5.9	1542	9	AG386981
c 8	58.8	5.9	996	9	CNS0039G
c 9	58.8	5.9	2015	9	AG311015
c 10	58.6	5.9	928	9	CNS0039G
c 11	58.4	5.8	1132	8	BZ695089
c 12	58.4	5.8	1335	9	CT646802
c 13	58	5.8	1101	9	CNS0039G
c 14	57.4	5.7	1025	9	CNS014J2
c 15	57.4	5.7	1201	9	CNS01522
c 16	57	5.7	1225	9	CNS0161D
c 17	56.8	5.7	1067	9	AG525799
c 18	56.6	5.7	774	8	BZ517772
c 19	56.4	5.6	815	9	AG056601
c 20	56.4	5.6	1202	8	CC262481
c 21	56.2	5.6	1092	9	CNS020K7
c 22	56	5.6	442	8	BZ518729
c 23	56	5.6	1200	9	CNS016CO
c 24	56	5.6	1758	9	CL509408

ALIGNMENTS

RESULT 1	CNS0039G/c	1101 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION	BAC08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL063921				
VERSION	AL063921.1	GI:4941778			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1. (bases 1 to 1101)				
TITLE	Genoscope.				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
FEATURES	Location/Qualifiers				
source	1. .1101				
	/organism="Drosophila melanogaster"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:7227"				
	/clone="BACR08K10"				
	/clone_lib="RPCI-98"				
	/note="end : TET3"				
ORIGIN					
Query Match	6.6%;	Score	66.2;	DB 9;	Length 1101;
Best Local Similarity	16.8%;	Pred. No.	0.0027;		


```

QY 560 CACTCGACTCGTGGTATTAAGTACTCGGTACGACCGCTGCAATTACTAATCGACC 619
DB 904 HMCNCHHMECTCHHHHTYHMYTCHHWHHWHHWHHWHHWHHWHHWHHWHHWHH 963
QY 620 TCGATTACATCACTTTTACGACTGCTTCAATGCTCAATCTTAATGAACGAGATTGG 679
DB 964 HMYHMMHMYCHYCTCHTHATTHYHMYCTCYHYCTTWTYTAHMYHMYHMYHMY 1023
QY 680 ACCCATACATAATATGACAAATGCTTCCAAAGAAACATGCTCTTATAGTGAATA 739
DB 1024 WWWHWHHWHHWHHWHHWHHWHHWHHWHHWHHWHHWHHWHHWHHWHHWHHWH 1083
QY 740 TC 741
DB 1084 HM 1085

RESULT 5
CNS020K7 1092 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 222L11 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL175696
VERSION AL175696.1 GI:7813753
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645
REFERENCE 2
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143
REFERENCE 3 (bases 1 to 1092)
GENOSCOPE.
Direct Submission
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
Location/Qualifiers
1..1092
/organism="Tetraodon nigroviridis"
/mol_type="Genomic DNA"
/db_xref="taxon:99883"
/clone_lib="222L11"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG222CF061P1-end : T7"
ORIGIN
Query Match 6.0%; Score 60.2; DB 9; Length 1092;
Best Local Similarity 38.7%; Pred No. 0.046; Mismatches 120; Indels 0; Gaps 0;
Matches 104; Conservative 45;

```

```

QY 656 AAATCTTAATGAAGCAGATTTTGACCCATACATAATATGACAAATGCTTCCAAAGA 715
DB 714 AAAAAAANWTAATWAAAWAAAAAANAANAANAANAANAANAANAANAANAANA 773
QY 716 AACATGGCTCTTATAGTGAATATCGTTAGACTCTTATAGAAAGATCTGAATTTATTA 775
DB 774 WAAAWWWTTTATTTTAAWWATWAWTTTATTTTAAAWAAANAANAANAANAANA 833
QY 776 TAAGATAGTGGTTTTTCTTTCTTTTCTTATCTAAGGAGTAAAGCAACCATGAATAGA 835
DB 834 AAAWAAATTAATTTTAAANAANAANAANAANAANAANAANAANAANAANAANA 893
QY 836 AAAGGCTTAGTAACTATATATCAAGGAATGCTGTTTTTCTTTAAATATGGAATAAAT 895
DB 894 WAAWTTWTTTCTTTTAAWTTTAAATWTTTCTTTTCTTTTCTTTTCTTTTCTTTT 953
QY 896 TTGTGATATAGAGATTAGATCAATTAA 924
DB 954 AAAAAAATATAAANAANAANAANAANAANAANAANAANAANAANAANAANA 982

RESULT 6
CNS00KAE 1101 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC:
DEFINITION BACR39P05 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL077628
VERSION AL077628.1 GI:4956903
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
GENOSCOPE.
Direct Submission
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org the BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department
of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain v2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="Genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR39P05"
/clone_lib="RPCI-98"
/note="end : TET3"
ORIGIN
Query Match 5.9%; Score 59.4; DB 9; Length 1101;
Best Local Similarity 37.5%; Pred. No. 0.066; Mismatches 323; Indels 1; Gaps 1;
Matches 209; Conservative 25;
2 ATATAGACTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 61

```

```

Db 40 ATATATAAATTANATATNTANANTATAAANTNTNTANANTNTNTANANTNTNTATATAAANAANTCTN 99
Qy 62 ATAATAAAGCTCAGATCTATATCAATTAACAGCAATCAAGGTCTACAGCGGTGAGAGA 121
Db 100 TCTATATAGATATATATATATNTNTNTANATATATATATNTANANTNTANANTNTATATATA 159
Qy 122 AGATTAATGATGATGATTCATTCATATATTCAGATCATTAATGATAGGGAATATCAAGC 181
Db 160 TNANTATANTANANANANANANATATATATANANTNTNTANTNTATAGKGNWNWATA 219
Qy 182 AATAAATAACGATAAATGCGATTAAGCAATTAAGATAAGGAGAAATGATTCACCCCAATATTGAAT 241
Db 220 TANATATANTNTNTNTNTNTNTATATATATATANNNANANATATANANANTNARATAN 279
Qy 242 GAGTGGATGATCTCTCTTTTGCAATGATGAATGATGGCAATATCTAGAAATGTTGGG 301
Db 280 TANANATAGWGATTTATTTATATAAAAAAATAATATATATATATATATATAAATTATAA 339
Qy 302 ACCCTCTCGGATCTAATGAAAAAGTAT-GGAATAGTAGATAATCGAATCTCTTTAGAA 360
Db 340 TAATTAATTAATTTAAAWAAWAATTTWTATTTATATNAAAAATTTATATATATAAAA 399
Qy 361 AGGTAGTGAATGCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAATATTTTATATCAG 420
Db 400 TAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 459
Qy 421 AGAATATATACATCCCTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 480
Db 460 ATAATATTTTATAWAAATTTAATAATATATATATATATATATATATATATATATATATA 519
Qy 481 TCAATCTTAAAGTACATACACCAAGAAATATTCATATAAATATTTTTCGAATTTCTAT 540
Db 520 TAAATTTTAAATTTAAWAAATTTAATAWAAATANAATAAATAAATAAATAAATAAATAA 579
Qy 541 TATAAATACTAGCTGTTA 558
Db 580 TATWAAAAAATAAATAA 597

-RESULT 7
AG386981/c
LOCUS Mus musculus molossinus DNA, clone:MSMg01-201G10.TJ, genomic survey
DEFINITION
ACCESSION AG386981
VERSION AG386981.1 GI:47998186
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
REFERENCE Hattori,M., Toyoda,A., Noguichi,H., Kojima,T. and Sakaki,Y.
AUTHORS BAC end Sequences of Library MSMg01
TITLE Unpublished
JOURNAL
REFERENCE Hattori,M., Toyoda,A., Noguichi,H., Kojima,T. and Sakaki,Y.
AUTHORS Direct Submission
TITLE Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
JOURNAL and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY

```

```

Vector : pBACE3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
Location/Qualifiers
source
1..1542
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-201G10.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 5.9%; Score 59; DB 9; Length 1542;
Best Local Similarity 43.8%; Pred. No. 0.079;
Matches 420; Conservative 0; Mismatches 525; Indels 15; Gaps 4;

Qy 6 AGACTTTTCATTAAATTTAATTTATGTCATCAGAAAGGATTAAACCTAGTTTAATGATAA 65
Db 1075 AAAATATATATTAATAATAAANTATTTAATAAATAATAAATAAATAAATAAATAAATAA 1016
Qy 66 TAACTTCAGATCTATAATCAATTAACAGCAATCAAGGTCTATAGCAGCGTTGAGAGAAGT 125
Db 1015 ATAAAAAAAATAATATTTAAAAATAAATAATAATAATAATAATAATAATAATAATAA 956
Qy 126 TAAATGTGATGYCATTCAATATTTCAAGATCATTAATGATAGGGGAATATCAAGCAATA 185
Db 955 ATAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 896
Qy 186 AATAACGATAAATGCGATTAAAGTAAATAAGGAGAAATGATTCACCCCAATATTGAATGAGG 245
Db 895 TAAATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 836
Qy 246 TGGATGATTCCTCTTTTTCAGCAATGATGAATGATGGGCAATCTAGATGATGTTGGGACCC 305
Db 835 AATATTATAAAAAATTAATTTAAATAATAATAAATAAATAAATAAATAAATAAATAAATAA 776
Qy 306 TTCTCGGATCTAATGAATAAAGTATGGAATAGTAGATATCGAATCTCTTTTAGAAGGTA 365
Db 775 TAAAAATATNAAAAATNAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 716
Qy 366 GTGATTGCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAATAATTTTATATCAGAGAT 425
Db 715 TAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 656
Qy 426 ATTAGATCCCTCTCTCCCTATCTCTTTTCTATTTATATGGGACATCTCCTCAATCAAT 485
Db 655 AATAAATAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 596
Qy 486 CCTAAAGTACATACACCAAGAATAATTCATAAATAAATAATTTTTCGAATATCTTATATAA 545
Db 595 ATTAAATNTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 542
Qy 546 AAATAGCTGTTAGCACTCGACCTCGTGGTATTTGACTACTCTCGGTAGGACCCCTGTCA 605
Db 541 AAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 482
Qy 606 TTTTACTAATCGACCTCGATTCATCATCTTTTACGATCTGCTTCTCATGCAATCTTAAAT 665
Db 481 TTTATATTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 422
Qy 666 GAAAGCAGATTTTGACCCCATACATAATATGACAAAAATTCGTCCTCAAGAAACATGGCT 725
Db 421 AAA---ATATATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 365
Qy 726 CTTATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGAATTTTATATATAGATAGT 785
Db 364 ATATTTTATTAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 305
Qy 786 GTTTTTTTCTCTTTT---CTTTTCATATCTAAGGAGTAAAGCAACCATGATGAATAAGAGCC 841

```


Query Match	5.9%;	Score 58.8;	DB 9;	Length 2015;
Best Local Similarity	42.0%;	Pred. No. 0.086;		
Matches 386;	Conservative 1;	Mismatches 532;	Indels 1;	Gaps 1;
Qy	8	ACITTTTGATTAAATTAATTAATATGTGATGACAAAGGATTAAACCTAGTTAAATGATTAATA	67	
Db	234	AATTATATTAAAAATAATAAAAAATTAATTAATAAATAATTAATAAATAAAAAAATAA	293	
Qy	68	ACTTCAGATCTATTAATCAATTAACAGCAATCAACGGTCATAGCAGCGTTGAGAGAGATTA	127	
Db	294	AATTAACCCAAAAAATAATAGATAAAAAATAAAAAAATAAAAAATAAAAAATAAAAAA	353	
Qy	128	AATGTGATGTYCAITTCATTAATTTCAAGATCATTAATGATGGGAATATCAAGCAATAAA	187	
Db	354	AATAATAAAAAAATAAAAAAGATATAACATAAACAACAAAAATAAATAAAAAATAA	413	
Qy	188	TACGATTAATGGCATTAAGATAAATAAGGAGATGATTCACCAATATTTGAATGAGGTG	247	
Db	414	AAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA	473	
Qy	248	GATGATTCCTCTTTTGACATGATGAATGATGGCAAACTAGAAATGTTGGGACCGCTT	307	
Db	474	AATAATATAATAAAACGGTAATAAATAATATAAATAAATAAATAAATAAATAATA	533	
Qy	308	CTCGGATCTAATGAAAAAGTATGGAATAGTAGATAATCGAATCTCTTTAGAAAGGTAGT	367	
Db	534	TTATTTTATAAACCCAAATATAATAAANGTAATAATAAAAAATAAGAAAAATAAATAA	593	
Qy	368	GATTGTCCTTTATCTAGAGAGAAAGTCTGCTTTTCAAAGAATAATTTTATCAGAGAATAT	427	
Db	594	TTAGATAAAAAAATAAATAATNAANAAGATAAATAAATAAATAAATAAATAAATAA	653	
Qy	428	TACATCCCGCTCTCCCTATCTCTTTTCTATTTATATGCGGACATTCCTCAATCAATCC	487	
Db	654	AAAAAATAAANAATAAATAATTTTTTTTTTATATTTTTTACAAAAAGAGNAATAAATAA	713	
Qy	488	TAAAGTACATACACCAAGAAATATCAATAAAATAATTTTTTGAATATCTTATTATAAAA	547	
Db	714	TAATAAAGATATAAGATAGATATCTCTATAAGTATCTAAGATAAATAATAAAGA	773	
Qy	548	ACTAGCTGTTAGCACTCGA CCGGTGCGYATTGCACTACTCGGTTAGAGCCCTGTCAATT	607	
Db	774	AAAAGTANCAACCCACCAACCTTTATTTTTTTATTTTCTCTCCCGCCCGCCCGCAC	833	
Qy	608	TACTAATCGACCTCGATTACATCACTTCTACGATCTGCTTCTATGTCATAATCTTAATGA	667	
Db	834	ACAAAAAATAAATAAATAAATAATACITTTCTTTTCAAAAAAATAAAAAAATAAAG	893	
Qy	668	AAGCAGATTTTGACCCCTACAATAATATGACAAATAATTGCTTCCAAAGAAACATGGCTCT	727	
Db	894	AAAAAAGAGGGGGGGTAATAAATTTTATAATATCTAGAATTTATAAAAAATAAAAAATAA	953	
Qy	728	TATAGTGAATATCGTTAGACTGTTATAGAAAGATCTGAATTTTATTATTAAGAAATAGTGT	787	
Db	954	TATTAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1013	
Qy	788	TTTTTTCTTTCTTTTCTATCTAAGGAGTAAGCAACCATGAAATGAAAGGGCTTAGTA	847	
Db	1014	AATAATAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1072	
Qy	848	ACTATATATCAAGGAATGGTGTGTTTTTCTTTTAAATATGGATAAAAAATTTGTGAATATAG	907	
Db	1073	AATACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1132	
Qy	908	AAGATTAGATCAATTAACAA	927	
Db	1133	AAATATAAAAAATAAATAAAAA	1152	

RESULT 10				
CNS00DKY				
LOCUS	CNS00DKY	928 bp	DNA	linear
DEFINITION	Driscophila melanogaster genome survey sequence T7 end of BAC #			GSS 04-JUN-1999

BACR27A24 of RPCI-98 library from *Drosophila melanogaster* (fruit fly), Genomic survey sequence.

AL071865
AL071865.1 GI:4948170
GSS.

Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; *Drosophila*.
1 (bases 1 to 928)
Genoscope.

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley *Drosophila* Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org/TheBDGP/Drosophila>

melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and 387 libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers
1..928
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR27A24"
/clone_lib="RPCI-98"
/note="end : 17"

ORIGIN

Query Match 5.9%; Score 58.6; DB 9; Length 928;
Best Local Similarity 32.1%; Pred. No. 0.097;
Matches 157; Conservative 67; Mismatches 258; Indels 7; Gaps 1;

QY 452 TTTTCTATTATATGGGACATTCCTCAATCATCTCTAAAGTACATACACCAAGATAT 511
Db 287 TTATTTTCAGAGACACACAGACATATTAGATCCAAAATATATACGTTATATAT 346
QY 512 TCAATAAAATATTTTTCGAATATCTTATTATAAAAAAGTACGTTGAGCACTCGACCTCG 571
Db 347 ATATATATATATATATATATATATKVINNNKKKGGGGGGGNNNNNGTGTGSG 406
QY 572 GTCGYATTGACTACTCGGTTAGACGCCCTGTCATTATCTAATCGACCTCGATTACATCA 631
Db 407 GGGCCNNNNCCNCCACANANANNNNNCCCCCCCCNNNNNNNNNNNNNNNNNN 466
QY 632 CTTTCTACGATACGTCTTCATGTCAAATCTTAATGAAGACAGATTTTGACCCATACAATA 691
Db 467 NNCNNNGGATTWTATATATATATAAATAATATGWTAAATTTTATKAWAAAAA 526
QY 692 ATATGACAAATTCGTTCCAAAGAAACATGCGCTCTTAT-----AGTGAATATCGTT 744
Db 527 AWAATTTTWTWTWTWTWAAAAAATATTTTWTWTWTWTWTWTWTWTWTWTWTWT 586
QY 745 AGACTGTATAGAAAGATCTGAATTTTATTATAGAAATAGTGTTTTCTTTCTTTTC 804
Db 587 TTATTTAAATTAATTTTATATWTAATAAATAATATWTAATAATATWTTWTWTA 646
QY 805 ATATCTAAGAGTAAAGCAACCATGAATAGAAAAGGCTTAGTAACATATATCAAGGAA 864
Db 647 WTWTATATAWAAAAWAAAAWAAAAWAAAAATTTTAAATTTTAAWAAAAWAAAAATTTA 706

(http://www.chori.org/bacpac/ordering_information.htm).

ORIGIN

Query Match 5.8%; Score 58.4; DB 9; Length 1335;
 Best Local Similarity 40.0%; Pred. No. 0.11;
 Matches 214; Conservative 1; Mismatches 319; Indels 1; Gaps 1;

QY 15 ATTAAATTAATAATTGTATGACAAAGGATTAAACCTAGTTAATGATATAATCTTCAG 74
 Db 1122 ANAAAAATTAATTTTTTATATAATAAAAAAATAAAAAAATAAAAAAATAAAAA 1063

QY 75 ATCTATATCAATTAACAGCAATCACGGTCATAGCAGCGTTGACAGAGATTAATGTA 134
 Db 1062 AAATANATTTTATATAAAAAAATAATAATTTTATATAAAAAAATAATAATNTNAA 1003

QY 135 TGTTCATTCAATATTTCAAGATCAATTAATGATGGGAATATCAAGCAATAATAAAGAT 194
 Db 1002 NNATATNAANNNTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATTT 943

QY 195 AAATGGCATTAAGTAATAAGGAGATGATTCACCAATATTCGAATGA-GGTGGATGAT 253
 Db 942 ATTTTAAATNNNAATTTTATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATNTN 883

QY 254 TCTTCTTTTTCACATGATGATGATGGCAAAATAGTATGTTGGACCTTCTCGGA 313
 Db 882 NATTAATTTTATAATAAATTTTANNNTATANAANAANNTTTTAAANNATATAAAA 823

QY 314 TCTAATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 373
 Db 822 TTTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 763

QY 374 CTTTTATCTAGAGAGAGTCTGCTTTTCAAGATATTTTATCAGAGATATTCATC 433
 Db 762 TATTTTTTAAAAAATAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 703

QY 434 CCCCTCTCCTCTATCTCTTTTCTATTATATGAGGACATCTCTCAATCAATCCTAAAG 493
 Db 702 TNNANNTATATATNT 643

QY 494 TACATACCAAGATATTCATTAATAATTTTTTTTGAATATCTTATATAAAAA 548
 Db 642 TATATAATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAATTTTAA 588

RESULT 13
 CNS003BD 1101 bp DNA linear GSS 03-JUN-1999
 Drosophila melanogaster genome survey sequence TET3 end of BAC #
 BACR08K08 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL064091
 VERSION AL064091.1 GI:4941847
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoe and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

Scori digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or fillers for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers

1..1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACH08K08"

/clone_lib="RPCI-98"

/note="end : TET3"

ORIGIN

Query Match 5.8%; Score 58; DB 9; Length 1101;
 Best Local Similarity 35.9%; Pred. No. 0.13;
 Matches 207; Conservative 58; Mismatches 312; Indels 0; Gaps 0;

QY 238 GAATGAGGTGATGATCTCTTTTGGACAAATGATGATGATGATGATGATGATGATG 297
 Db 1087 GGATAGGTGATGATCTCTTTTGGACAAATGATGATGATGATGATGATGATGATG 1028

QY 298 TGGGACCCCTTCGGATCTAATGAAAGATGATGATGATGATGATGATGATGATG 357
 Db 1027 RWTTRTWTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTW 968

QY 358 GAAAGGTAGTATGATCTCTTTTATCTAGAGAGAAAGTCTGCTTTTCAAGAAATTTTAT 417
 Db 967 TATTTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNT 908

QY 418 CAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 477
 Db 907 WTTATWTTATWTTATWTTATWTTATWTTATWTTATWTTATWTTATWTTATWTT 848

QY 478 CAATCAATCTTAAAGTACATACACCAAGATATCAATAAATATTTTGAATATTC 537
 Db 847 TWTANATWTTTAAATWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNTWNT 788

QY 538 TATATAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 597
 Db 787 WATWATATTTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 728

QY 598 CCTGTCTATCTAATCGACCTGATTCATCCTCTACGATCTCTCTCATGCTCA 657
 Db 727 TTTANWNTTATTAATTTTTTTTATATTTTTTTTTTTTTTTTTTTTTTTTATTTT 668

QY 658 ATCTTAATGAAAGCAGATTTTGACCCATACATAAATATGACAAATTTGCTTCCAAAGAA 717
 Db 667 ATAAAAATTTTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 608

QY 718 ACATGCTCTTATAGTGAATATCGTTAGATGTTTATAGAAAGATCTGATTTATTTATA 777
 Db 607 AAATKWTTTATKATAA 548

QY 778 AGAATAGTCTTTTTTCTTTCTTTTCTATCTATCTAAGG 814
 Db 547 AAATTATTAAGG 511

RESULT 14
 CNS014J2 1025 bp DNA linear GSS 26-JUL-1999
 Drosophila melanogaster genome survey sequence T7 end of BAC
 BACH11L1 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL104216
 VERSION AL104216.1 GI:5615827
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Search completed: November 11, 2004, 12:05:44
Job time : 3112.21 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C 2	1	60.4	6.0	1141	4	US-09-806-708B-22
	2	52	5.2	1141	4	US-09-806-708B-22
	3	50.4	5.0	1055	4	US-09-806-708B-23
C 5	4	49.2	4.9	5360	4	US-10-204-708-66
	5	48.8	4.9	832	4	US-09-621-976-2813
	6	47.2	4.7	1055	4	US-09-806-708B-23
C 7	7	46.4	4.6	1122	4	US-09-248-796A-729
	8	46	4.6	1140	4	US-09-270-767-29120
	9	46	4.6	2396	4	US-09-270-767-13204
C 10	10	45.8	4.6	7218	1	US-08-232-463-14
	11	44.6	4.5	319608	4	US-09-539-333D-1
	12	44.6	4.5	319608	4	US-09-679-409-1
C 14	13	43.8	4.4	4285	3	US-09-410-464-1
	14	43.6	4.4	392	4	US-09-270-767-29271
	15	43.6	4.4	2031	4	US-09-270-767-13327
C 17	16	43.4	4.3	6113	4	US-10-204-708-13
	17	43	4.3	933	4	US-09-248-796A-9931
	18	42.6	4.3	3967	4	US-09-366-715-5
C 20	19	42.2	4.2	19224	2	US-08-487-828B-13
	20	42	4.2	539	4	US-09-270-767-25193
	21	42	4.2	643	4	US-09-270-767-30522
C 22	22	42	4.2	1188	4	US-09-270-767-9895
	23	42	4.2	1842	4	US-09-270-767-14367
	24	41.8	4.2	1316	4	US-09-270-767-2481
C 25	25	41.8	4.2	1316	4	US-09-270-767-17763
	26	41.8	4.2	640681	4	US-09-790-988-1
	27	41.6	4.2	2238	4	US-09-792-024-34


```

Db 63 TKWGTTCWAGTWRWAAKKTKKWCSTAMNNWTTTCTARKWRTGTRWTKTNNATGT 122
Qy 133 GATGTCATTCATATTTCAAGATCATTAATGATAGGGAATATCAAGCAATATAAAG 192
Db 123 RWTGYMTNNNGSTWTAARYKTRWYCYTAMWYGASWAGNASTRTTYTWKWKMK 182
Qy 193 ATAAATGGCAATTAAGTAATAAGGAGAATGATTCAACCAATATTGAATGAGGTGGATGA 252
Db 183 RKSAPAAATGRARYMRAWTWAPRTGWTKAAYAATWNNNNNNNAKAAKCAATTWGR 242
Qy 253 TTCTTTCTTTTGCAATGATGATGGCAAAATGATGAATGTTGGGACCCCTTCCTCGG 312
Db 243 AKSNCTCTTAGTTTCTATCCWATTCGAGWATKKWTKTKTSAAMGTMWNNNNNNNTTK 302
Qy 313 ATCTAATGAAAAGATAGTATGATAGTATGATATCGATCTCTTTAGAAAGGTAGTGTG 372
Db 303 AAMTAARAWWWSATTTWAAATTSRTWYGRKTAANNNGTTCWTRWATWKKWK 362
Qy 373 TCTTTTATCTAGAGAGAAAGTCTCTTTTCAAGAAATATTATTATCAGAGAATATTACAT 432
Db 363 TKGTTWNNNGRRTYCTGTTTCKMATTTTAKANNCTTAAGWKCTTMMNTTAAKATTYAT 422
Qy 433 CCCCTCTCTCCCTATCTCTTTTCTATTATATATGAGACATTCTCTCAATCAATCCTAAA 492
Db 423 CYKSMWNGT-SYRYAAARYTWYAWTR- - - -RYAYANNNTKTWKAATWTKKCCCT 477
Qy 493 GTACATACCAAGAAATATTCAATAAATATTTTTTGAATATCTTATTATAAATACTAG 552
Db 478 ANNTAANYTKSSANCTSRTRWTKNCWRAGSKTASMGRAVARAYWTKWNTAWYCW 537
Qy 553 CTGTTAGACTCGACCTCGGTGVTATGACTACTCGGTTACGAGCCCTCTCTATTACTA 612
Db 538 WYVYRAGAAMTAMWMTSATCYATAATTAGTCAGAGGSTAKGNNNNNNNNCCATCAR 597
Qy 613 ATCGACCTCGATTACATCACTCTTCTACGATACCTCTCATGTCAAATCTTAATGAAGCA 672
Db 598 WKCTAASACMANATTCYCYAANNATYWANATCGWNAKTATATWNNNNNNNGTW 657
Qy 673 GATTTTGACCATACAAATAATATGACAAATTTGTTCCAAAGAAACATCGCTCTTATAG 732
Db 658 TNNNNNAKVASATWYAAAMTAATKYARTANTAMAGAYARAAAYTTRTANNGACTTT 717
Qy 733 TGAATATCTGTAGCTGTTATAGAAAGATCTGNAATTTTATTAAGAATAGTGTTTT 792
Db 718 TTTNNNTGGRNTTAAARGWANNNNNNNNNNNNNNGACWARTTTTATANCSTNNNNNN 777
Qy 793 TCTTTCTTTTTCATCTAAGGAGTAAAGCAACCATGAATAGAAAAGCGCTTAGTAATAT 852
Db 778 NAYATTTNTATTTTWTTRKANNNNNNNNNNNNAAYYGAAAKNNNTTMCWTKAWKAWATGA 837
Qy 853 ATATCAAGGAATGGGTGTTTTCTTTTAAATATGGATA 890
Db 838 ATTTNAGTATATNNNNNNATATTTTYYKAATNGKACTA 875

RESULT 4
US-10-204-708-66
; Sequence 66, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEX, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058-8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173-8

```

```

; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529-7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826-1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 66
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-66

Query Match 4.9%; Score 49.2; DB 4; Length 5360;
Best Local Similarity 52.4%; Pred. No. 0.014; 98; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 0;

Qy 684 ATACAATAATATGACAAAAATTCCTTCCAAAGAAACATCGGCTCTTATAGTGAATATCGT 743
Db 1793 AAAAAAAGATAAAAAATTTGGATATAGTAGTATTGGTTATTATTATTAAAGATGATA 1852
Qy 744 TAGACTGTTATAGAAGATCTCAATTTATTTATAAGATAGTGTGTTTTTCTTTCTTTT 803
Db 1853 AAATGTTTTTTTGGAAATGAAAAAGTTTTTATGTAAGATATATTTTTTATTATTAGTTA 1912
Qy 804 CATATCTAAGGAGTAAAGCAACCATCAATAGAAAAAGGCTTAGTAACTATATATCAAGGA 863
Db 1913 GTTTTAAAAACGTTAGGAAGAGTATATATTAATGAAGTATTTTTTATGGTAATATTA 1972
Qy 864 ATGGTGTGTTTTTCTTTAAATATGGAT 889
Db 1973 AAGTTGTTTTTATTTTGTAAATTCGAAT 1998

RESULT 5
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match 4.9%; Score 48.8; DB 4; Length 832;
Best Local Similarity 10.8%; Pred. No. 0.01;
Matches 29; Conservative 136; Mismatches 103; Indels 0; Gaps 0;

Qy 721 TGGCTCTTATAGTGAATATCGTTAGACTGTTATAGAAGATCTGAATTTATTATAAGA 780
Db 1 YAWYWKYTTTAKCWTWKWSWSYWKYTKYTWWRKWKKAWKWKYKWTWYVW 60
Qy 781 ATAGTGTGTTTTTCTTTTCTATCATCTAAGCAACCAACCATGAATAGAAAAG 840
Db 61 RYAMGTYKKKAMCRTKTKKKKKGYMMWYMGWRSYAMWTRTWGTGYVYRSWYVWR 120
Qy 841 CTTAGTAATATATCAAGGAATGTTGTTTCTTTTAAATATCGATAAAAAATTGTG 900
Db 121 YRCWKKAAYRKTYTCYSSKGMWTKWKRKCAWTTWWKKTYWAAATRYWMMCMWTKRWRAS 180

```

```
QY 901 AATATAGAGATTGATCAATTAACAAAGGTTATCGTGGAGTGAAGCAGAGCGGACC 960
Db 181 WYCWGKRWKSWKRSYSARSARCCYCSGMSGMSKWKYWRWRWGRWATGAGM 240
QY 961 TATGTGTTATAGTAGAGGGTCAACCACT 988
Db 241 KAWRASOMRRKYAGKSKTSYKSMWMCW 268

RESULT 6
US-09-806-708B-23/c
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806.708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 2001-04-03
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 23
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1055)
; OTHER INFORMATION: consensus sequence of A.t. and L.a. FAE1 promoters
US-09-806-708B-23

Query Match 4.7%; Score 47.2; DB 4; Length 1055;
Best Local Similarity 18.8%; Pred. No. 0.027;
Matches 90; Conservative 145; Mismatches 242; Indels 1; Gaps 1;

QY 284 AATACTAGATGTTGGACCCCTCTCGGATCTTAATGAAAAAGTATGGAATAGATA 343
Db 529 ANMRCARWTRTYCKSTAMCTYGNAMWYAAYSAGNTSSMARWTTANNATAGGYMRA 470
QY 344 ATCGAATCTTTAGAGAGTAGTGTCTTTTATCTAGAGAGAGCTGCTTTTCA 403
Db 469 AWAGTWMAAMNNTRTRYAWTRWARYTTTIRSAKNXKSWRGATWRAATMTTAA 410
QY 404 AAGATATTTTATCAGAGAATATTACATCCCTCTCTCCTATCTCTTTTCTATTTA 463
Db 409 NKAGAMWMTTAAAGNNTWAAAATKMAAWCARAYCCNNWAAACMAKWKMAWTWYA 350
QY 464 TATGGGACATTCCTCAATCAATCCTAAAGTACATACACCAAGAAATATTCAATAAATAT 523
Db 349 WGAACNNNNKAMYCRRAWMYSAWTTTWAAATSWWKWYTTTRKTTMAAAANNNNNW 290
QY 524 TTTTGTGAATATCTATTATAAAACTAGCTGTGTAGCACTCGACCTCGGTCGYTATTGAC 583
Db 289 AKCKTTSAMWMMWATWCT-CGARTWGAATYMAAACCTAAGAGNSMTYWCWATYNGTT 231
QY 584 TACTCGGTAGAGCCCTGTCATTACTAATCGACCTCGATTACATCATCTTCTACGATA 643
Db 230 MTNNNNNNKAWTTRTKTMAWCAYTTWTRWYKYTYCYATYTYTSMYMGKWWYWARAA 171
QY 644 CTGCTTCATGTCATATCTTAATGAAGCAGATTTTGGACCCATCAATAATATGACAAAT 703
Db 170 YYASTNCTWSTCRWKTARGWYIYMYRYTWAKSGCANNKWRCAWYVYCATNNNAW 111
QY 704 TGCTTCCAAAGAAACATGCGTCTTATAGTGAATATCGTTAGACTGTTATAGAAAGA 761
Db 110 YACAYWYWTAGKAAWNNKNTASGKWWYAMMKTWYWAYTWCAACWATAKRRTAAM 53

RESULT 7
US-09-248-796A-729
```

```
; Sequence 729, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074.725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096.409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 729
; LENGTH: 1122
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-729

Query Match 4.6%; Score 46.4; DB 4; Length 1122;
Best Local Similarity 50.9%; Pred. No. 0.042;
Matches 110; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 1 AATATAGACTTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 60
Db 863 AATGGAAGCTGTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 922
QY 61 GATAATACTTCAGATCTATATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
Db 923 CACCACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 982
QY 121 AAGATTAATGATGTYCAITCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
Db 983 ATATTGATACCTTTACAAAAATCAAAATTAATTAATTAATTAATTAATTAATTAATTA 1042
QY 181 CAATAATAACCATTAATGCGCATTAATTAATTAATTAATTAATTAATTAATTAATTA 216
Db 1043 CTTTACGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1078

RESULT 8
US-09-270-767-29120
; Sequence 29120, Application US/09270767
; Patent No. 67034591
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 29120
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-29120

Query Match 4.6%; Score 46; DB 4; Length 1140;
Best Local Similarity 50.5%; Pred. No. 0.053;
Matches 109; Conservative 1; Mismatches 106; Indels 0; Gaps 0;

QY 18 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 77
Db 182 ACATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 241
QY 78 TATAATCAATTAACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 137
Db 242 GAAAAAATCTTTATGAAAACTTAATAAAGGAAAACTGAAAGAAAAATCGAGGCATATAT 301
QY 138 YCATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 197
```

Db 302 CTATTGAAGCTCTTTAATTATTGAAAGCTGCAAACTCGGGATCACCAGTGGATAA 361
QY 198 TGGATTAAAGTAATAAAGAGAAATGATTCACCAA 233
Db 362 TCGATTAAAGTACATAAATAAATTAAATAAAACCA 397

RESULT 9
US-09-270-767-13204
; Sequence 13204, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13204
; LENGTH: 2396
; TYPE: DNA
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-13204

Query Match 4.6%; Score 46; DB 4; Length 2396;
Best Local Similarity 50.5%; Pred. No. 0.066;
Matches 109; Conservative 1; Mismatches 106; Indels 0; Gaps 0;

QY 18 AAATTAATTAATTGTTGATGCAAAAGGATTAACCTAGTTAATGATTAATTAATTCAGATC 77
Db 182 AACATAAATTGTAGTTCTTTGAAAGAAATTAACCACTGGAATTTAGAAAATAAAAC 241
QY 78 TATAATCAATTAACAGCAATCACGGTTCATAGCAGCGTTGAGAGAAAGATTAAATGTGATG 137
Db 242 GAAAAATCTTTGAAAACTTAATTAAGAAACTGAAAGAAATCGAGGCATATAT 301
QY 138 YCATTCATATTTCAAGATCATTAATGATAGGGGAATATCAAGCAATAAATAACGATAA 197
Db 302 CTATTGAAGCTCTTTAATTATTATTGAAAGCTGCAAACTCGGGATCACCAGTGGATAA 361
QY 198 TGGCATTAAAGTAATAAGGAGATGATTCACCAA 233
Db 362 TCGATTAAAGTACATAAATAAATTAAATAAAACCA 397

RESULT 10
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:

; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 INMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 4.6%; Score 45.8; DB 1; Length 7218;
Best Local Similarity 3.7%; Pred. No. 0.1;
Matches 14; Conservative 210; Mismatches 157; Indels 0; Gaps 0;

QY 12 TTGATTAAATTAATTAATTTGATGACAAAGGATTAACCTAGTTAATGATAATAACTT 71
Db 1459 TTAAGAGATAGAGAATTTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1400
QY 72 CAGATCTAATCAATTAACAGCAATCAGGTCATAGCAGCGTTGAGAGAGAAATTAATG 131
Db 1399 RRR 1340
QY 132 TCGATGYCATTCATTAATTTCAAGATCAATTAATGATAGGGGAATATCAAGCAATAAATAC 191
Db 1339 RRR 1280
QY 192 GATAAATGCATTAAAGTAATAAGAGAAATGATTCACCAATATTGAATGAGTGGATG 251
Db 1279 RRR 1220
QY 252 ATTCTCTTTTGACAATGATCAATGATGGCAATATAGAAATCTGGACCCCTCTCG 311
Db 1219 RRR 1160
QY 312 GATCTAATGAAAAAGATATGGAATAGTAGAATAATCGAATCTCTTTAGAAAGTAGTGTG 371
Db 1159 RRR 1100
QY 372 GTCTTTTCTAGAGAGAAAG 392
Db 1099 RRR 1079

RESULT 11
US-09-539-333D-1
; Sequence 1, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bouguetret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET.047AUS
; CURRENT APPLICATION NUMBER: US/09/539,333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971

```

; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..1107
; OTHER INFORMATION: 5'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1108..1289
; OTHER INFORMATION: exon A g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14877..14920
; OTHER INFORMATION: exon B g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 18778..18862
; OTHER INFORMATION: exon Bbis g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25593..25740
; OTHER INFORMATION: exon C g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29388..29502
; OTHER INFORMATION: exon D g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29967..30282
; OTHER INFORMATION: exon E g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 34666..64812
; OTHER INFORMATION: exon F g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 65505..65853
; OTHER INFORMATION: exon G g35018 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 65854..67854
; OTHER INFORMATION: 3'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 94124..94964
; OTHER INFORMATION: exon g35017
; FEATURE:
; NAME/KEY: exon
; LOCATION: 201188..201234
; OTHER INFORMATION: exon S g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 214676..214793
; OTHER INFORMATION: exon T g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215702..215746
; OTHER INFORMATION: exon U g35030 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 216836..216915
; OTHER INFORMATION: exon V g35030 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 213818..215818
; OTHER INFORMATION: 3'regulatory region g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215819..215941
; OTHER INFORMATION: exon R complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 215819..215975
; OTHER INFORMATION: exon Rbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 216661..216952
; OTHER INFORMATION: exon Qbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 217027..217061
; OTHER INFORMATION: exon Q1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 229647..229742
; OTHER INFORMATION: exon X complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 230408..230721
; OTHER INFORMATION: exon P complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231272..231412
; OTHER INFORMATION: exon Obis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231787..231880
; OTHER INFORMATION: exon O2 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 231870..231879
; OTHER INFORMATION: exon O1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 234174..234321
; OTHER INFORMATION: exon O complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 237406..237428
; OTHER INFORMATION: exon Nbis complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 239719..239807
; OTHER INFORMATION: exon N2 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 239719..239853
; OTHER INFORMATION: exon N complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240528..240569
; OTHER INFORMATION: exon M117 complement g34872 gene
; FEATURE:
```

NAME/KEY: exon
LOCATION: 240528..240596
OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240617
OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240644
OTHER INFORMATION: exon M52 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240824
OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240994
OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..241685
OTHER INFORMATION: exon M1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240800..240993
OTHER INFORMATION: exon M51 complement g34872 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 241686..243685
OTHER INFORMATION: 5'regulatory region g34872 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 290652..292652
OTHER INFORMATION: 3'regulatory region g34665 gene
FEATURE:
NAME/KEY: exon
LOCATION: 292653..292841

Query Match 4.58; Score 44.6; DB 4; Length 319608;
Best Local Similarity 49.84; Pred. No. 0.6; Indels 0; Gaps 0;
Matches 113; Conservative 0; Mismatches 114

Qy 117 AGAAGATTAATGATGTCATTCATTAATTTTCAAGATCATTAATGATAGGGGAATAT 176
Db 104809 ACATAATTTTAAATTAGAAAATTTATGACAAAACCTCGTTTCATTCCTTTTGGAAAA 104868
Qy 177 CAACCAATAAATACGATAAATGCGCATTAAGTAATTAAGAGATGATTCACCCCAATAT 236
Db 104869 CGATACTAACCACTAGTAAATGTTAAAGTCCTGCAATATGCACAGGGGTTAATAAAATTT 104928
Qy 237 TGAATGAGGTGATGATTCCTTTTGGACAATGATGATGCGCAATATCTAGAAATG 296
Db 104929 AGGATGACATGGCTCAGGCTTGTTTAGACAAATAGACACTGATAGGCTGTTCATAAAC 104988
Qy 297 TTGGACCTTCCTCGATCTAATGAAAAAGATGGAATAGTAGATA 343
Db 104989 TTGTTATATTCCTTGGAGAAATGAGGCCCATATATAATGAATATA 105035

RESULT 12

US-09-679-409-1
Sequence 1, Application US/09679409
Patent No. 655316
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueret, Lydie
APPLICANT: Essieux, Laurent
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS
FILE REFERENCE: 53.US15.CIP
CURRENT APPLICATION NUMBER: US/09/679,409

CURRENT FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 09/539,333
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 09/416,384
PRIOR FILING DATE: 1999-10-12
PRIOR APPLICATION NUMBER: 60/168,088
PRIOR FILING DATE: 1999-11-30
NUMBER OF SEQ ID NOS: 134
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 319608
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 199122..201122
OTHER INFORMATION: 5'regulatory region
NAME/KEY: exon
LOCATION: 201123..201234
OTHER INFORMATION: exon S
NAME/KEY: exon
LOCATION: 201123..201560
OTHER INFORMATION: exon S2
NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U
NAME/KEY: exon
LOCATION: 216836..216994
OTHER INFORMATION: exon V
NAME/KEY: exon
LOCATION: 216836..217077
OTHER INFORMATION: exon V2
NAME/KEY: exon
LOCATION: 217671..217764
OTHER INFORMATION: exon V1
NAME/KEY: exon
LOCATION: 227655..227736
OTHER INFORMATION: exon V4
NAME/KEY: exon
LOCATION: 238715..238919
OTHER INFORMATION: exon V3
NAME/KEY: exon
LOCATION: 240440..240673
OTHER INFORMATION: exon W
NAME/KEY: exon
LOCATION: 240440..241153
OTHER INFORMATION: exon W2
NAME/KEY: exon
LOCATION: 241072..241291
OTHER INFORMATION: exon X
NAME/KEY: exon
LOCATION: 244353..244561
OTHER INFORMATION: exon Y
NAME/KEY: exon
LOCATION: 246273..247802
OTHER INFORMATION: exon Z
NAME/KEY: misc feature
LOCATION: 247803..249803
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 8316
OTHER INFORMATION: 99-27943-150 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 21672
OTHER INFORMATION: 99-27935-193 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 65485
OTHER INFORMATION: 8-128-33 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 95396

```
OTHER INFORMATION: 99-31960-363 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 107281
OTHER INFORMATION: 99-24656-260 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 160640
OTHER INFORMATION: 99-24639-163 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 160876
OTHER INFORMATION: 99-24634-108 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 168974
OTHER INFORMATION: 99-7652-162 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 170810
OTHER INFORMATION: 99-16100-147 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 173358
OTHER INFORMATION: 99-5862-167 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 189957
OTHER INFORMATION: 99-5919-215 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 197163
OTHER INFORMATION: 99-24658-410 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 200778
OTHER INFORMATION: 8-303-235 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 202651
OTHER INFORMATION: 8-300-221 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 202679
OTHER INFORMATION: 8-300-193 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 203378
OTHER INFORMATION: 8-299-128 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 204138
OTHER INFORMATION: 8-296-213 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 204605
OTHER INFORMATION: 8-252-190 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 204934
OTHER INFORMATION: 99-24644-194 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 205206
OTHER INFORMATION: 8-295-248 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 205329
OTHER INFORMATION: 8-295-125 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 206064
OTHER INFORMATION: 8-293-130 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 206545
OTHER INFORMATION: 8-292-198 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 207313
OTHER INFORMATION: 8-251-322 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 208285
OTHER INFORMATION: 8-289-322 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 208960
OTHER INFORMATION: 8-287-249 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 209423
OTHER INFORMATION: 8-287-86 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 209631
OTHER INFORMATION: 8-285-319 : polymorphic base A or G
```

```
NAME/KEY: allele
LOCATION: 210361
OTHER INFORMATION: 8-283-278 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210463
OTHER INFORMATION: 8-283-176 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210486
OTHER INFORMATION: 8-283-153 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210583
OTHER INFORMATION: 8-283-56 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 210879
OTHER INFORMATION: 8-282-345 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210964
OTHER INFORMATION: 8-282-260 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 210979
OTHER INFORMATION: 8-282-245 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 211050
OTHER INFORMATION: 8-282-174 : variable motif AAAGG or GAAGGAAGGAAGGAAGA
NAME/KEY: allele
LOCATION: 211132
OTHER INFORMATION: 8-282-92 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 211247
OTHER INFORMATION: 8-281-367 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211315
OTHER INFORMATION: 8-281-299 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211366
OTHER INFORMATION: 8-281-248 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 212520
OTHER INFORMATION: 8-279-197 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 212821
OTHER INFORMATION: 8-278-289 : polymorphic base C or T
NAME/KEY: allele
```

```
Query Match 4.5%; Score 44.6; DB 4; Length 319608;
Best Local Similarity 49.8%; Pred. No. 0.6;
Matches 113; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 117 AGAGAGATTAAATGATGATGATTCATTCATTAATTTCAAGATCATTAATGATAGGGGAATAT 176
Db 104809 ACATAATTTTAAATTTAGAAATTTAGCAAAATCGTTCATTCCTTTTGTGAAAAA 104868
QY 177 CAAGCAATAATAACGATAAATGGCATTAAAGTAATAAGGAGATGATTCACCAATAT 236
Db 104869 CGTACTTAACCACTAGTAAATGTTAAGTCCTGATTATGCACAGGGTTAATAAATTT 104928
QY 237 TGAATGAGGTGATGATTTCTTTTGCATGATGATGATGCGCAAACTAGATG 296
Db 104929 AGGATGACATGGCTCAGGCTTTTAAAGACAAATGACACTGATAGGCTGTTCCATAAAC 104988
QY 297 TTGGGACCCCTTCGATCTAATGAAAAAGTATGGAATAGTAGATA 343
Db 104989 TTGTTATATTGCTTGGAGAAATGAGGCCCATATATATATGATATATA 105035
```

```
RESULT 13
US-09-410-464-1
; Sequence 1, Application US/09410464
; Patent No. 635892
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: poplar and other plant species.
```

```

; FILE REFERENCE: 53375
; CURRENT APPLICATION NUMBER: US/09/410,464
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: 09/287,700
; EARLIER FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/080,851
; EARLIER FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4285
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
US-09-410-464-1

Query Match          4.4%; Score 43.8; DB 3; Length 4285;
Best Local Similarity 51.9%; Pred. No. 0.26;
Matches 122; Conservative 0; Mismatches 112; Indels 1; Gaps 1;

QY 684 ATACAATAATGACAAAATGCTTCCAAAGAA-AACATGGCTCTTATAGTGAATATCG 742
DB 1378 ATCAAGGATAAATTCAAAATTCGATCCAAATAACACATTAAGTAAAGGACTTATGG 1437
QY 743 TTGACTGTTATAGAAAGATCTGAATTTATTATAGAAATAGTGTCTTTTCTTTCTTT 802
DB 1438 TATGATGGATCATATAATATAAGGCTGAATTAACAACATTTTTTTTATTAGATCTGT 1497
QY 803 TCATATCTAAGGAGTAAGCAACCATGAATAGAAAAGGCTTAGTAATATATATCAAAAG 862
DB 1498 TTATTTTACGTTTAAAAATATTTTGAATTTATTTTATTTTATATAAATTAATAT 1557
QY 863 AATGGCTTTTCTTTAAATATCGATAAAAAATTTGTAATAGATAGATATGAT 917
DB 1558 TTTTAGATCATTTTAATACGTTAATATAAAAAATAATTTTTTTAAAAAATTTAT 1612

RESULT 14
US-09-270-767-29271/c
; Sequence 29271, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 29271
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-29271

Query Match          4.4%; Score 43.6; DB 4; Length 392;
Best Local Similarity 50.5%; Pred. No. 0.15;
Matches 103; Conservative 1; Mismatches 100; Indels 0; Gaps 0;

QY 11 TTTGATTAAATTAATTAATTTATGATGACAAAGGATTAACCTAGTTTAATGATAATACT 70
DB 247 TATTATTACTCTACACGGAGAATGAACCGGAATTAACCTAATTAATAGTCAGCGAAACG 188
QY 71 TCAGATCTATAATCAATTAACAGCAATCACGGTCTATAGCAGCGTTGAGAGAAGATTAAAT 130
DB 187 TTAGCAAGAATATCAGAAAGAGGAAAGCAGTACACAAATGTAATACGAACACATATGAT 128
QY 131 GTGATGTYCATTCATATTTCAAGATCATTAATGATAGGGGAATATCAAGCAATAATAA 190
DB 127 AAAATTACCATATAATATTTTATGCAAGATAAATTTCTAGTTAAGAAGCAACAAAGTAAA 68

191 CGATAAATGGCATTAAAGTAAATA 214
67 TAATAAATAAGTTTGAAAAACA 44

Search completed: November 11, 2004, 12:09:12
Job time : 109.152 secs
```

```

RESULT 15
US-09-270-767-13327/c
; Sequence 13327, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 13327
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13327

Query Match          4.4%; Score 43.6; DB 4; Length 2031;
Best Local Similarity 50.5%; Pred. No. 0.24;
Matches 103; Conservative 1; Mismatches 100; Indels 0; Gaps 0;

QY 11 TTTGATTAAATTAATTAATTTATGATGACAAAGGATTAACCTAGTTTAATGATAATACT 70
DB 247 TATTATTACTCTACACGGAGAATGAACCGGAATTAACCTAATTAATAGTCAGCGAAACG 188
QY 71 TCAGATCTATAATCAATTAACAGCAATCACGGTCTATAGCAGCGTTGAGAGAAGATTAAAT 130
DB 187 TTAGCAAGAATATCAGAAAGAGGAAAGCAGTACACAAATGTAATACGAACACATATGAT 128
QY 131 GTGATGTYCATTCATATTTCAAGATCATTAATGATAGGGGAATATCAAGCAATAATAA 190
DB 127 AAAATTACCATATAATATTTTATGCAAGATAAATTTCTAGTTAAGAAGCAACAAAGTAAA 68

191 CGATAAATGGCATTAAAGTAAATA 214
67 TAATAAATAAGTTTGAAAAACA 44

Search completed: November 11, 2004, 12:09:12
Job time : 109.152 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2004, 10:27:15 ; Search time 535.761 Seconds
(without alignments)
10091.039 Million cell updates/sec

Title: US-10-009-966C-2_COPY_2155_3155
Perfect score: 1001
Sequence: 1 aataagacttttgattaaa.....accactactagaatccgg 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3625171 seqs, 2700493622 residues

Total number of hits satisfying chosen parameters: 7250342

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	55.6	5.6	158001	16	US-10-211-179-11
c 2	54.8	5.5	37515	17	US-10-433-793-28
c 3	54.4	5.4	113515	15	US-10-311-455-2148
4	52.6	5.3	8201	15	US-10-240-485-64
5	52	5.2	13584	15	US-10-311-455-588
c 6	51.4	5.1	27890	17	US-10-741-601-5686
7	51.4	5.1	126872	17	US-10-741-601-5738
8	51.4	5.1	3673778	15	US-10-312-841-2
9	50	5.0	5413	16	US-10-221-714A-418
c 10	50	5.0	11745	15	US-10-240-453-206
11	49.4	4.9	11694	16	US-10-221-714A-422
c 12	49.4	4.9	3673778	15	US-10-312-841-2

13	49.2	4.9	5360	15	US-10-204-708-66	Sequence 66, Appl
14	49.2	4.9	5360	15	US-10-311-455-1910	Sequence 1910, Ap
15	49.2	4.9	5360	16	US-10-240-589C-106	Sequence 106, App
c 16	49.2	4.9	5743	15	US-10-311-455-2041	Sequence 2041, Ap
17	49.2	4.9	7306	15	US-10-311-455-1609	Sequence 1609, Ap
c 18	49.2	4.9	34588	17	US-10-433-793-90	Sequence 90, Appl
19	49	4.9	9810	15	US-10-311-455-400	Sequence 400, App
20	48.6	4.9	6109	15	US-10-311-455-299	Sequence 299, App
c 21	48.8	4.9	6109	16	US-10-221-613-33	Sequence 33, Appl
22	48.8	4.9	34722	17	US-10-322-281-700	Sequence 700, App
c 23	48.4	4.8	12507	15	US-10-311-455-271	Sequence 271, Appl
24	48.4	4.8	37515	17	US-10-433-793-27	Sequence 1757, Ap
c 25	48.2	4.8	7461	15	US-10-311-455-1757	Sequence 1757, Ap
26	47.8	4.8	15732	14	US-10-239-676-95	Sequence 95, Appl
27	47.8	4.8	15732	15	US-10-240-453-107	Sequence 107, App
c 28	47.8	4.8	48509	17	US-10-322-281-510	Sequence 510, App
29	47.8	4.8	3673778	15	US-10-312-841-1	Sequence 1, Appli
30	47.2	4.7	1223197	13	US-10-027-632-179264	Sequence 179264,
c 31	47.2	4.7	1223197	15	US-10-027-632-179264	Sequence 179264,
32	47	4.7	5572	15	US-10-311-455-1397	Sequence 1397, Ap
33	47	4.7	5572	15	US-10-240-452-59	Sequence 59, Appl
c 34	47	4.7	5979	14	US-10-239-676-18	Sequence 18, Appl
35	47	4.7	5979	15	US-10-240-453-28	Sequence 28, Appl
36	47	4.7	11422	15	US-10-311-455-192	Sequence 192, App
37	47	4.7	11422	16	US-10-257-166-18	Sequence 18, Appl
c 38	46.8	4.7	14316	16	US-10-221-613-408	Sequence 408, App
c 39	46.6	4.7	40862	15	US-10-311-455-2046	Sequence 2046, Ap
40	46	4.6	13328	17	US-10-433-793-104	Sequence 104, App
41	46	4.6	20579	17	US-10-433-793-104	Sequence 104, App
c 42	45.8	4.6	5184	16	US-10-221-613-383	Sequence 383, App
43	45.8	4.6	16545	15	US-10-311-455-23	Sequence 23, Appl
c 44	45.8	4.6	122614	13	US-10-087-192-1726	Sequence 1726, Ap
45	45.6	4.6	6240	15	US-10-311-455-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1

US-10-211-179-11

- ; GENERAL INFORMATION:
- ; APPLICANT: Nicholas M. Dean
- ; APPLICANT: Kenneth W. Dobie
- ; FILE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOTRYSYL PHOSPHATASE ACTIVATOR EXPRI
- ; TITLE REFERENCE: PTS-0011
- ; CURRENT APPLICATION NUMBER: US/10/211,179
- ; CURRENT FILING DATE: 2002-08-01
- ; NUMBER OF SEQ ID NOS: 119
- ; OTHER INFORMATION: n = a, t, c, or g
- US-10-211-179-11

Query Match 5.6%; Score 55.6; DB 16; Length 158001;
Best Local Similarity 42.6%; Pred No. 0.67;
Matches 346; Conservative 1; Mismatches 460; Indels 5; Gaps 1;

QY	117	AGAGAAGATTAAATGATGATTCATTCATTAATTTCAAGATCATTAAATGATGGGAATAT	176
DB	22124	ATATATAATAA	22183
QY	177	CAAGCAATATAACGATTAATGGCAATTAAGTAATAGGGAATGATTCACCAATAT	236
DB	22184	TAATATAATATAATATAATATAATATAATATAATATAATATAATATAATATAATAT	22243
QY	237	TGAATGAGGTGATGATTCCTCTTTTGCAATGATGATGATGGCAATAGTACTAGATG	296
DB	22244	ATAATATAAAATAATATAATATAATATAATATAATATAATATAATATAATATAATAT	22303
QY	297	TTGGGACCCCTTCGGAICTAATGAAAAGATGGAATAGTAGATATCGAATCTCTTT	356
DB	22304	ATATAATATAATATAATATAATATAATATAATATAATATAATATAATATAATATAAT	22363
QY	357	AGAAAGGTAGTGTCTCTTTTATCTAGAGAAAGTCTGCTTTTCAAGATATATTTTA	416

RESULT 10
US-10-240-453-206/c
; Sequence 206, Application US/10240453
; Publication No. US20030148326A1

```

;
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;
; FEATURE:
;
; NAME/KEY: unsure
;
; LOCATION: (9105)
;
US-10-240-453-206

```

Qy	13	TGATTAAATTAAATTAATTTGTATGACAAAGGATTAAACCTAGTTAATGATAATAACTTC	72
Db	6199	TTAACAAATCTTTAAATCATAATTAATAAAAAATCGAAAAACCAATAAAAAAAA	6140
Qy	73	AGATCTATAATCAATTTAAACAGCAANTCACGGTCATAGCAGC-GTTGAGAGAAGATTAAATG	131

DB	6139	ATATAACTATTAAATTTTAAAAAATCAATATAAAAAATATCTCTAAAAAATTAATAAAAATAAT	608
QY	132	TGATGTCATTCAATATTTCAAGATCATTAATGATAGGGGAATATCAAGCAATAAATAAC	191
DB	6079	TATATATCATCTAAAAAATAATTTAAAAATATATACAATCAATTAATTAATTAAT	6020
QY	192	GATAAATGGCATTAAATTAAGGAGAGATGATTCACCCCAATATTGAATGAGGTGGATG	251
DB	6019	ATCTACAATAAATACTATATCTTTAATAATAAAAAATTAATAAAAAATTAATAATTA	5960
QY	252	ATCTCTCTTTTGACAATGATGAATGATGGGCAAAATCTAGAAATGTTGGACCCCTCTCG	311
DB	5959	AACTCTCTAATTTCTATAACTAAAAATCTAAAAAATCAAAAAATTAATAAATAATTA	5900
QY	312	GATCTAATGA--AAAAAGTATGGAATAGTAGAATAATCGAATCTCTTTAGAAAGGTAGTGA	359
DB	5899	CATTTAATAATTAATAAATCTTTTAAATAATCAATAATTAATAAATCTATAAAAAATTAATAAT	5840

QY 370 TTGCTTTTCTAGAGAAAGTCTGCTTTTCAAGAAATATTTTATCAGAGAAATATTA 429
| | | | |
Db 5839 TATCTATAAAAAATAATAATAAAACTTACAAAAATCCATAATTAATAAAAAATTATTA 5780
| | | | |
QY 430 CA 431
| | | | |
Db 5779 AA 5778

RESULT 11

US-10-221-714A-422
; Sequence 422, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 422
; LENGTH: 11694
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-422

Query Match 4.9%; Score 49.4; DB 16; Length 11694;
Best Local Similarity 51.3%; Pred. No. 4.8;
Matches 139; Conservative 0; Mismatches 131; Indels 1; Gaps 1;
QY 12 TTGATTAAATTAATTAATTTGTATGACAAAGGATTAAACCTAGT-TAATGATAATAACT 70
| | | | |
Db 9647 TTGCGTAAAGATTAAATTTTATTAATAGGAAATTAATTAATTTATGTTAATTTACG 9706
| | | | |
QY 71 TCAGATCTATATCAATTAACAGCAATCACCGTCATAGCAGCGTTGAGAGAAATTAAT 130
| | | | |
Db 9707 TGATAAATAAAGTAGATTTATATATATAGAGGTGGAAAGAGATATTTAAGAGATAATAAAA 9766
| | | | |
QY 131 GTGATGTCATTCATATTTCAAGTCAATTAATGATGGGAATCAAGCAATAAATAA 190
| | | | |
Db 9767 ATGAATAAGTTTGAATTTTAAATTAATTTATTAAGAGGAAATTTAGAAATAATAA 9826
| | | | |
QY 191 CGATAAAATGGCATTAAAGTAATAAGGAGATGATTCACCAATTAATGAATGAGGTGGAT 250
| | | | |
Db 9827 ATGGATGGGAATTTAGAAATTAAGTCTTAATAGTAATTTATTTAGTAGGATTGTGAAG 9886
| | | | |
QY 251 GAATCTCTTTTGGACATGATGATGAG 281
| | | | |
Db 9887 GGTTTTATTTTATTTATTTATTTTGTG 9917
| | | | |

RESULT 12

US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1

GENERAL INFORMATION:

; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1206/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match 4.9%; Score 49.4; DB 15; Length 3673778;

Best Local Similarity 47.3%; Pred. No. 50;
Matches 181; Conservative 0; Mismatches 201; Indels 1; Gaps 1;
QY 140 ATTCAATATTTCAGATCATTAATGATAGGGGAATATCAAGCAATATAAATAACGATAATG 199
| | | | |
Db 574390 ATTTAATATTCTATATCTCAAAATCCAAATATAAAAAATAAAAAATAAAAACTAAAAA 574331
| | | | |
QY 200 GCATTAAAGTAAATAGGAGAAATGATTCACCCCAATATTGAATGAGGTGGATCTTCT 259
| | | | |
Db 574330 AAAATTATCAATAAATAATAATTTAAAAAATCTCCAAATTTAAACAAAAACATAAACCTAAT 574271
| | | | |
QY 260 TTTTGACATGATGATGATGGCAATACCTAGAAATGTTGGACCCCTTCTCGGATCTAAT 319
| | | | |
Db 574270 AATTAAAAAATAAATAAACCACCAATATAAACAACCAACCAACCAACCAACCAATC 574211
| | | | |
QY 320 GAAAAAAGTATGGAATAGTAGAATAATCGAATCTCTTTAGAAAGGTAGTGATGCTTTTA 379
| | | | |
Db 574210 ATAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAACA 574151
| | | | |
QY 380 TCTAGAGAGAAATGCTGCTTTTCAAGAGATATTTTATCAGAGAAATATTACATCCCCCTC 439
| | | | |
Db 574150 TCTAACAAATAAAAAATAAACAACCAAAATTTTCATCAAAACTATATAAACCCGAAA 574091
| | | | |
QY 440 TCTCCCTATCTCTTTTCTTATTTATATGGACATTCCTCAATCAATCTATAAAGTACATA 499
| | | | |
Db 574090 AAAATCACCACCTTTTTCCAATATAAAAAATAAAAAATAAATAAATAAATAAATAAATA 574031
| | | | |
QY 500 -CACCAAGAAATATTCATAAAT 521
| | | | |
Db 574030 TCACCAAAAAATATCCAAAAAAT 574008
| | | | |

RESULT 13

US-10-204-708-66
; Sequence 66, Application US/10204708
; Publication No. US20030141852A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1

```
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 66
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-66

Query Match          4.9%; Score 49.2; DB 15; Length 5360;
Best Local Similarity 52.4%; Pred. No. 3.8;
Matches 108; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 684 ATACAATAATATGACAAAATTCCTCCAAAGAAAACATGGCTCTTATAGTGAATATCGT 743
Db 1793 AAAAAAAGATAAATAATTTGGATATAGTAGTATGGTTATTTATTAAGATGATA 1852

QY 744 TAGACTGTTATAGAAAGATCTGAATTTATTTAAGAATAGTGTCTTTCTTTCTTTT 803
Db 1853 AAATGTTTTTTTGGAAATGAAAAGTTTTATGTAAGATATATTTTATTTATTTAGTTTA 1912

QY 804 CATATCTAAGGAGTAAGCAACCATCAATAGAAAGGCTTAGTAACTATATATCAAGGA 863
Db 1973 AGTTGTTTTATTTTGTAAATTTGGAAT 1998

RESULT 14
US-10-311-455-1910
; Sequence 1910, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting Cytosine Methylation
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1910
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1910

Query Match          4.9%; Score 49.2; DB 15; Length 5360;
Best Local Similarity 52.4%; Pred. No. 3.8;
Matches 108; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 684 ATACAATAATATGACAAAATTCCTCCAAAGAAAACATGGCTCTTATAGTGAATATCGT 743
Db 1793 AAAAAAAGATAAATAATTTGGATATAGTAGTATGGTTATTTATTAAGATGATA 1852

QY 744 TAGACTGTTATAGAAAGATCTGAATTTATTTAAGAATAGTGTCTTTCTTTCTTTT 803
Db 1853 AAATGTTTTTTTGGAAATGAAAAGTTTTATGTAAGATATATTTTATTTATTTAGTTTA 1912

QY 804 CATATCTAAGGAGTAAGCAACCATCAATAGAAAGGCTTAGTAACTATATATCAAGGA 863
Db 1973 AGTTGTTTTATTTTGTAAATTTGGAAT 1998

RESULT 15
US-10-240-589C-106
; Sequence 106, Application US/10240589C
; Publication No. US20040076956A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting Cytosine Methylation
; TITLE OF INVENTION: DNA repair
; FILE REFERENCE: 5013.1008
; CURRENT APPLICATION NUMBER: US/10/240.589C
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03972
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 148
; SEQ ID NO 106
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-589C-106

Query Match          4.9%; Score 49.2; DB 16; Length 5360;
Best Local Similarity 52.4%; Pred. No. 3.8;
Matches 108; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 684 ATACAATAATATGACAAAATTCCTCCAAAGAAAACATGGCTCTTATAGTGAATATCGT 743
Db 1793 AAAAAAAGATAAATAATTTGGATATAGTAGTATGGTTATTTATTAAGATGATA 1852

QY 744 TAGACTGTTATAGAAAGATCTGAATTTATTTAAGAATAGTGTCTTTCTTTCTTTT 803
Db 1853 AAATGTTTTTTTGGAAATGAAAAGTTTTATGTAAGATATATTTTATTTATTTAGTTTA 1912

QY 804 CATATCTAAGGAGTAAGCAACCATCAATAGAAAGGCTTAGTAACTATATATCAAGGA 863
Db 1973 AGTTGTTTTATTTTGTAAATTTGGAAT 1998

Search completed: November 11, 2004, 14:31:46
Job time : 548.761 secs
```

This Page Blank (uspto)